GULF OF ALASKA PACIFIC OCEAN PERCH: STOCK ASSESSMENT, SURVEY DESIGN AND SAMPLING

A

THESIS

Presented to the Faculty of
the University of Alaska Fairbanks

In Partial Fulfillment of the Requirements
For the Degree of

DOCTOR OF PHILOSOPHY

By

Dana Henry Hanselman, B.S., M.S.
Juneau, Alaska

May 2004

GULF OF ALASKA PACIFIC OCEAN PERCH: STOCK ASSESSMENT,

SURVEY DESIGN AND SAMPLING

By

Dana Henry Hanselman

RECOMMENDED:

Advisory Committee Chair

Director, Fisheries Division

APPROVED:

Dean, School of Fisheries and Ocean Sciences

Dean of the Graduate School

May 7, 2004

Date

Abstract

Pacific ocean perch (*Sebastes alutus*) stock size in the Gulf of Alaska has been difficult to assess because of an imprecise survey biomass index. This imprecision has been attributed to low sampling effort on a species with an aggregated distribution. In this thesis, I examined the importance of estimated survey biomass in the stock assessment and ways to improve them. First, I presented the complete stock assessment for 2003, with an analysis of uncertainty. Uncertain parameters included natural mortality, recruitment, and biomass estimates. Second, I examined adaptive cluster sampling (ACS) as a method to reduce survey uncertainty. ACS results provided lower estimates of mean abundance and lower standard errors than did simple random sampling (SRS). Bootstrapping suggested that the ACS mean may be a superior measure of central tendency. ACS results were better than SRS, but not as dramatically as suggested by previous literature. I used simulations to explore why ACS did not perform optimally. These simulations showed that it would be necessary to sample over 10% of the population to obtain large gains in precision. This is impractical for a large marine population. I explored the use of hydroacoustic data recorded on survey vessels to gain precision in biomass estimation. I used the data to (1) develop a catch prediction model based on near-bottom backscatter, (2) simulate an adaptive design, (3) apply ratio estimation in double sampling using hydroacoustic data, and (4) post-stratify survey data. Using hydroacoustic data in these designs showed gains in precision over SRS and may be useful. Finally, I used the *S. alutus* age structured model presented above to simulate effects of five factors: survey measurement error, catchability trends, a second biomass index, data source weighting, and sensitivity of prior distributions. Simulations showed that the stock assessment model was ineffective at high measurement error and was unable to detect trends in the data. A second biomass index yielded gains in model precision. The weight given lengths measured in the fishery was most important because of its long time series, and the prior distribution on natural mortality was most influential because it was difficult to estimate.

Table of Contents

List of Figures

List of Tables

List of Other Materials

Acknowledgments

While this thesis has my name on the title page, I am by no means the sole contributor. So many people have contributed to this thesis and my university experience, and this is my opportunity to express my gratitude.

Foremost, I want to thank my major professor, Terry Quinn, for calling me unexpectedly that day in July of 1998, when I was working at a gym, considering what to do with my life. He offered me a fine opportunity to further my education in a new and interesting place. Since that fateful day, I have completed my M.S. in Juneau and now my doctorate degree under his advisement. He has always been there when I needed his expertise and advice, but has always let me feel free to pursue my own research interests. For all that he has done for me, I am deeply grateful. Terry also is second author on Chapters 2-5. For Chapter 2, Terry was integral in fine-tuning the statistics and sampling designs that I developed. In Chapter 3, he provided some of the code required to run the simulations. For Chapters 4 and 5, Terry was instrumental in refining some of the statistics and interpretation of results. Terry also helped me revise much of the writing herein.

Second, I want to thank the Trusty Three, that is, the other three venerable gentlemen on my committee, Drs. Tony Gharrett, Lewis Haldorson, and Jon Heifetz, for all their helpful comments and discussions and their willingness to be on my committee not once, but twice. Jon is the coauthor of two of the manuscripts included in this thesis. He was the senior author of the previous Pacific ocean perch stock assessments and developed the early framework from which Chapter 1 was developed. Dr. Heifetz also helped with the sampling design in Chapter 2.

Third, I must thank the Auke Bay Laboratory staff for their support of my projects and advice including Dave Clausen, Chris Lunsford, Dr. Jeff Fujioka, Phil Rigby and Dean Courtney. Dave is a coauthor on Chapter 1, where he provided some of the historical Pacific ocean perch biology and fishery information. For Chapter 2, Dave and

Chris were critical contributors in setting up vessel contracts and logistics for the experiment.

Fourth, administrative staff in the Juneau office, Gabbi, Louisa and Deb: thank you for always making life's little daily challenges easy to overcome. Also, I want to thank, Dr. Bill Smoker who, as fisheries director, provided some excellent writing guidance.

This publication is the result of research sponsored by the National Marine Fisheries Service and Alaska Sea Grant with funds from the National Oceanic and Atmospheric Administration Office of Sea Grant, Department of Commerce, under grant no. NA 86RG0050 (project no.E/142-02), and from the University of Alaska with funds appropriated by the state. The U.S. Government is authorized to produce and distribute reprints for government purposes, notwithstanding any copyright notation that may appear hereon. Further support was provided by the Auke Bay Lab, Alaska Fisheries Science Center, National Marine Fisheries Service.

Dedication

I dedicate this thesis to my parents, Marge and Dave Hanselman, for always assuring me that an education was more important than money (an obvious prerequisite for a fisheries degree) and offering their perpetual support and love. I also dedicate this thesis to Dr. Kalei Shotwell, whose support, vision and wit has been with me every step of the way.

General Introduction

The assessment of fish populations is one of the most important and problematic areas in fisheries management. Assessing rockfish species in the Gulf of Alaska and elsewhere has been particularly difficult. At least 96 species of rockfish (genus *Sebastes*) inhabit the northern Pacific Ocean (Love et al. 2002). Many of the more abundant rockfish that inhabit the continental slope share several important features related to stock assessment, survey design, and sampling. These features include a prominent swim bladder, deep demersal existence, and a patchy population distribution. Capturing rockfish from great depth induces certain mortality from barotrauma to the swim bladder, which makes mark-recapture studies difficult. The aggregated distribution of some of the major commercial slope rockfish makes them difficult to sample precisely using conventional designs (Hanselman et al. 2001). Rockfish are slow-growing and long-lived with sporadic recruitment events which make them particularly susceptible to overfishing. I illustrate some of these problems and some potential solutions with the Pacific ocean perch (*Sebastes alutus*).

Pacific ocean perch is the dominant commercial rockfish species in Alaska (Hanselman et al. 2003). It is very long-lived (>80 years) and is highly aggregated, which makes the harvest of this species particularly difficult to manage. An accurate survey biomass estimate for Pacific ocean perch is needed to properly assess the status of the stock. The multispecies groundfish trawl surveys conducted by the National Marine Fisheries Service (NMFS) have not adequately sampled Pacific ocean perch, because they are highly aggregated. This poor sampling is a result of both the small area inhabited by Pacific ocean perch along the slope of the continental shelf as compared to more uniformly distributed flatfish found on the shelf itself, and because much of its primary bottom habitat is inaccessible to standard survey gear (Lunsford 1999). Large fluctuations in survey biomass estimates with great uncertainty occur as a result. Because of the aforementioned life history characteristics of Pacific ocean perch and their low mortality and moderate fecundity, actual rapid upward and downward changes in biomass in the absence of heavy exploitation are unlikely.

In this thesis, I examine how these sampling problems could be addressed in terms of survey and sampling designs, and what role these survey biomass estimates play in the resulting stock assessment.

In Chapter 1, I present the Stock Assessment and Fishery Evaluation (SAFE) document for Pacific ocean perch in the Gulf of Alaska. This manuscript is presented first because it provides a historical and biological background for Pacific ocean perch. Chapter 1 also shows how these biomass estimates are used in concert with a number of other data sources. I also explore model uncertainties (Patterson et al. 2001) and apply the Markov Chain Monte Carlo method of Bayesian integration (Gelman et al. 1995).

In Chapter 2, I present a field study using adaptive cluster sampling (Thompson 1990) to improve the precision of biomass estimates. The design has received much attention in the literature over the last decade as a technique to gain precision in abundance estimation for rare or clustered populations. This chapter discusses the sampling design used, the results, and some simulations of those results to explore the performance of the design on Pacific ocean perch and two other important rockfish species, shortraker (*S. borealis*) and rougheye (*S. aleutianus*) rockfish.

In Chapter 3, I present a simulation study exploring situations in which adaptive cluster sampling would be effective and situations in which it may not be. I compare a simulated population from the literature (Su and Quinn 2003), for which there large gains in precision by use of the adaptive design, to a population that has the characteristics of a population of *S. alutus*.

In Chapter 4, I present some applications of the use of hydroacoustic data recorded on survey vessels to enhance or supplement existing surveys. Hydroacoustic information has been used to identify fish populations since the 1920's (Kimura 1929) and has become important in many of the world's major fisheries (McLennan and Simmonds 1992). These data can be collected with little cost and effort and may be used as an approach to gain survey precision. In this chapter, I explore the vagaries of hydroacoustic data, specifically for deep, near-bottom fishes with uncalibrated survey echosounders. I then develop a model to use the raw backscatter information for three sampling designs: adaptive cluster sampling, double sampling, and post-stratification.

In Chapter 5, I present five simulation experiments that explore some of the important characteristics of 'Integrated Analysis' (Punt et al. 2001) stock assessments like the one conducted for Pacific ocean perch. I examine the effects of measurement error on survey biomass estimates, temporal trends in catchability, the addition of a secondary biomass index (like hydroacoustics), the effects of deliberate weighting of multiple data sources and the effects of using informative prior distributions on key model parameters. With this chapter, I put in perspective the relative importance of survey biomass estimates and other data sources in the overall stock assessment process.

References

Gelman, A., J.B. Carlin, H.S. Stern and D.B. Rubin. 1995. Bayesian data analysis. Chapman and Hall, London. 526 pp.

Hanselman, D.H., T.J. Quinn II, C. Lunsford, J. Heifetz and D.M. Clausen. 2001. Spatial implications of adaptive cluster sampling on Gulf of Alaska rockfish. In Proceedings of the 17th Lowell-Wakefield Symposium: Spatial Processes and Management of Marine Populations, pp. 303-325. Univ. Alaska Sea Grant Program, Fairbanks, AK.

Hanselman, D., J. Heifetz, J. Fujioka, and J. Ianelli. 2003. Gulf of Alaska Pacific ocean perch. In Stock Assessment and fishery evaluation report for the groundfish resources of the Gulf of Alaska. North Pacific Fisheries Management Council, Anchorage. pp. 429-479.

Kimura, K. 1929. On the detection of fish groups by an acoustic method. J. Imp. Fish Inst., Tokyo 24:41-5.

Love, M. S., M. Yoklavich and L. Thorsteinson. 2002. The Rockfishes of the Northeast Pacific. University of California Press, Berkeley, California.

Lunsford, C. 1999. Distribution patterns and reproductive aspects of Pacific ocean perch (*Sebastes alutus*) in the Gulf of Alaska. M.S. thesis. University of Alaska Fairbanks, Juneau Center, School of Fisheries and Ocean Sciences.

MacLennan, D.N. and J.E. Simmonds. 1992. Fisheries Acoustics. Chapman and Hall, New York.

Patterson, K., R. Cook, C. Darby, S. Gavaris, L. Kell, P. Lewy, B. Mesnil, A. Punt, V. Restrepo, D.W. Skagen and G. Stefansson. 2001. Estimating uncertainty in fish stock assessment and forecasting. Fish and Fisheries. 2: 125-157.

Punt, A.E., D.C. Smith, R.B. Thomson, M. Haddon, X. He, and J.M. Lyle. 2001. Stock assessment of the blue grenadier Macruronus novaezelandiae resource off south-eastern Australia. Mar. Freshwater Res. 52: 701-717.

Su, Z. and T. J. Quinn.II 2003. Estimator bias and efficiency for adaptive cluster sampling with order statistics and a stopping rule. Environmental and Ecological Statistics 10:17-41.

Thompson, S.K. 1990. Adaptive cluster sampling. Journal of the American Statistical Association 412:1050–1059.

1 Stock Assessment of Gulf of Alaska Pacific ocean perch[1]

1.1 Introduction

Pacific ocean perch (POP), *Sebastes alutus*, is the dominant fish in the slope rockfish assemblage and has been extensively fished along its North American range since 1940 (Westrheim et al. 1970). The species has a wide geographic range in the North Pacific from California, to the Bering Sea and Southwest to the Kuril Islands. Pacific ocean perch are viviparous, with internal fertilization, and release of live young. Spawning takes place in relatively deep water (>250m) in early winter. Fertilization takes place after the sperm is held in the female for a short time, followed by several months of gestation. Larvae are released in April-May. This parturition time corresponds with the large plankton blooms that occur in the spring in the Gulf of Alaska. Dependency on the timing of this bloom could be a reason for their sporadic recruitment. Identification of the larvae of POP is difficult and infrequent (Gharrett et al. 2001). Consequently there is considerable uncertainty about the early life history of the species. POP larvae are hypothesized to stay at depth of release for extended periods, and then move to shallower waters over several months. Larvae feed on varying sizes of copepods and larvae as they grow. During this stage, larvae are pelagic and do not settle into demersal existence for 2-3 years (Gunderson 1977, Haldorson and Love 1991). Among rockfish, POP juveniles have one of the lower daily growth rates of rockfish juveniles. Upon recruitment, juveniles settle on hard low-relief sediments (Love et al. 1991). Older fish are generally found between 150-350 meters in the summer time and deeper in the winter (Love et al. 2002). Pacific ocean perch are very slow growing and long lived with

---

[1] Hanselman, D., J. Heifetz, J. Fujioka, and J. Ianelli. 2003. Gulf of Alaska Pacific ocean perch. In Stock Assessment and Fishery Evaluation Report for the Groundfish Resources of the Gulf of Alaska. North Pacific Fisheries Management Council, Anchorage. pp. 429-479. Some formatting has changed, and a few generic sections common to all SAFE documents were removed.

natural mortality rates of about 0.05. Maximum age has been estimated to exceed 90 years (Leaman 1991). However, 90% of maximum size (~48 cm) is usually reached by 20-25 years of age.

Few studies have been conducted on the stock structure of Pacific ocean perch. Based on allozyme variation, Seeb and Gunderson (1988) concluded that Pacific ocean perch are genetically quite similar throughout their range, and genetic exchange may be the result of dispersion at early life stages. In contrast, preliminary analysis of mitochondrial DNA structure suggests that genetically distinct populations of Pacific ocean perch exist (A. J. Gharrett pers. commun., University of Alaska Fairbanks, October 2000). Withler et al. 2001 found genetically distinct populations on a small scale in British Columbia. Currently, genetic studies are underway that should clarify the genetic stock structure of Pacific ocean perch.

In 1991, the North Pacific Fisheries Management Council (NPFMC) divided the slope assemblage in the Gulf of Alaska into three management subgroups: Pacific ocean perch, shortraker/rougheye rockfish, and all other species of slope rockfish. In 1993, a fourth management subgroup, northern rockfish, was also created. These subgroups were established to protect Pacific ocean perch, shortraker/rougheye, and northern rockfish (the four most sought-after commercial species in the assemblage) from possible overfishing. Each subgroup is now assigned an individual ABC (acceptable biological catch) and TAC (total allowable catch), whereas prior to 1991, an ABC and TAC was assigned to the entire assemblage. Each subgroup ABC and TAC is further divided among three management areas of the Gulf of Alaska (Western, Central, and Eastern) based on distribution of exploitable biomass.

Amendment 41, which took effect in 1998, prohibited trawling in the Eastern area east of 140 degrees W. longitude. Since most slope rockfish, especially Pacific ocean perch, are caught exclusively with trawl gear, this amendment could have concentrated fishing effort for slope rockfish in the Eastern area to the relatively small area between 140 degrees and 147 degrees W. longitude that remained open to trawling. To ensure that such a geographic over-concentration of harvest would not occur, since 1999 the

NPFMC has divided the Eastern area into two smaller management areas: West Yakutat (area between 147 and 140 degrees W. longitude) and East Yakutat/Southeast Outside (area east of 140 degrees W. longitude). Separate ABC's and TAC's are now assigned to each of these smaller areas for Pacific ocean perch.

1.1.1 Fishery

1.1.1.1 Historical Background

A Pacific ocean perch trawl fishery by the U.S.S.R. and Japan began in the Gulf of Alaska in the early 1960's. This fishery developed rapidly, with massive efforts by the Soviet and Japanese fleets. Catches peaked in 1965, when a total of nearly 350,000 metric tons (mt) was caught. This apparent overfishing resulted in a precipitous decline in catches in the late 1960's. Catches continued to decline in the 1970's, and by 1978 catches were only 8,000 mt (Figure 1.1a). Foreign fishing dominated the fishery from 1977 to 1984, and catches generally declined during this period. Most of the catch was taken by Japan (Carlson et al. 1986). Catches reached a minimum in 1985, after foreign trawling in the Gulf of Alaska was prohibited.

The domestic fishery first became important in 1985 and expanded each year until 1991 (Figure 1.1b). Much of the expansion of the domestic fishery was apparently related to increasing annual quotas; quotas increased from 3,702 mt in 1986 to 20,000 mt in 1989. In the years 1991-95, overall catches of slope rockfish diminished because of the more restrictive management policies enacted during this period. The restrictions included: (1) establishment of the management subgroups, which limited harvest of the more desired species; (2) reducing levels of total allowable catch (TAC) to promote rebuilding of Pacific ocean perch stocks; and (3) conservative in-season management practices in which fisheries were sometimes closed even though substantial unharvested TAC remained. These closures were necessary because, given the large fishing power of the rockfish trawl fleet, there was substantial risk of exceeding the TAC if the fishery were to remain open. Since 1996, catches of Pacific ocean perch have increased again, as good recruitment and increasing biomass for this species have resulted in larger TAC's. In the last several years, the TAC's for Pacific ocean perch have been fully taken (or nearly so) in each management area except Southeastern. (The prohibition of trawling in Southeastern during these years has resulted in very little catch of Pacific ocean perch in this area.)

Detailed catch information for Pacific ocean perch in the years since 1977 is listed in Table 1.1a for the commercial fishery and in Table 1.1b for research cruises. The reader is cautioned that actual catches of Pacific ocean perch in the commercial fishery are only shown for 1988-2002; for previous years, the catches listed are for the Pacific ocean perch complex (a former management grouping consisting of Pacific ocean perch and four other rockfish species), Pacific ocean perch alone, or all *Sebastes* rockfish, depending upon the year (see Footnote in Table 1.1). Pacific ocean perch make up the majority of catches from this complex. The acceptable biological catches and quotas in Table 1.1 are Gulfwide values, but in actual practice the NPFMC has divided these into separate, annual apportionments for each of the three regulatory areas of the Gulf of Alaska. (As explained in the last paragraph of section 1.2, the Eastern area for Pacific ocean perch has been subdivided into two areas, so there is now a total of four regulatory areas for these two management groups.)

Historically, bottom trawls have accounted for nearly all the commercial harvest of Pacific ocean perch. In recent years, however, a sizable portion of the Pacific ocean perch catch has been taken by pelagic trawls. The percentage of the Pacific ocean perch Gulfwide catch taken in pelagic trawls increased from 2-8% during 1990-95 to 14-20% during 1996-98. In the years 1999-2002, the amount caught in pelagic trawls has remained moderately high, with annual percentages of 17.6, 10.3, 11.7 and 11.0, respectively.

Before 1996, most of the Pacific ocean perch trawl catch (>90%) was taken by large factory-trawlers that processed the fish at sea. A significant change occurred in 1996, however, when smaller shore-based trawlers began taking a sizeable portion of the catch in the Central area for delivery to processing plants in Kodiak. The following table shows the percent of the total catch of Pacific ocean perch in the Central area that shore-based trawlers have taken since 1996[2]:

1996	1997	1998	1999	2000	2001	2002
49	28	32	41	52	43	58

Factory trawlers continued to take nearly all the catch in the Western and Eastern areas.

1.1.1.2 Bycatch

Ackley and Heifetz (2001) examined bycatch in Pacific ocean perch fisheries of the Gulf of Alaska by using data from the observer program for the years 1993-95. For hauls targeting Pacific ocean perch, the major bycatch species were arrowtooth flounder, shortraker/rougheye rockfish, sablefish, and "other slope rockfish". (This was based only on data for 1995, because there was no directed fishery for Pacific ocean perch in 1993-94). More recent data (Gaichas and Ianelli summaries of NMFS Observer data) from 1997-2002 show that the largest bycatch groups in the combined rockfish trawl fishery are arrowtooth flounder, Pacific cod, and sablefish in that order. The same data set shows that the only major non-rockfish fishery that catches substantial Pacific ocean perch is the rex sole fishery, averaging 280 mt per year. Small amounts of Pacific ocean perch are also taken in other flatfish, Pacific cod, and sablefish fisheries (Gaichas and Ianelli summaries of NMFS Observer data).

1.1.1.3 Discards

Gulfwide discard rates[2] (% discarded) for Pacific ocean perch in the commercial fishery for 1991-2002 are listed as follows:

Year	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002
%Discard	18.4	29.4	79.2	59.7	19.7	17.2	14.5	14.0	13.8	11.3	8.6	7.2

The high discard rates for Pacific ocean perch in 1993 and 1994 can be attributed to its "bycatch only" status for most of this time period. Since then, discard rates for Pacific ocean perch have steadily decreased.

---

2 National Marine Fisheries Service, Alaska Region, Fishery Management Section, P.O. Box 21668, Juneau, AK 99802-1688. Data are from weekly production and observer reports through October 28, 2003.

1.2 Data

1.2.1 Fishery Data

1.2.1.1 Catch

Catches range between 2500 mt and 350,000 mt in the years 1961 to 2003. Detailed catch information for Pacific ocean perch is listed in Table 1.1a and shown graphically in Figure 1.1.

1.2.1.2 Age and Size composition

Observers aboard fishing vessels and at onshore processing facilities have provided data on size and age composition of the commercial catch of Pacific ocean perch. Ages were determined from the break-and-burn method (Chilton and Beamish 1982). Table 1.2 summarizes the length compositions from 1977-2003 (with several gaps). Table 1.3 summarizes age compositions from 1998-2002 for the fishery. Figures 1.3 and 1.4 show the distributions. The age compositions in all five years of the fishery data show strong 1987 and 1988 year classes. These year classes were also strong in age compositions from the 1996 and 1999 trawl surveys. The 1993 and previous surveys show more strength in the 1986 year class. The fishery age data shows high correlation when lagged, indicating ages and collections are consistent.

1.2.2 Survey Data

1.2.2.1 Biomass Estimates from Trawl Surveys

Bottom trawl surveys were conducted on a triennial basis in the Gulf of Alaska in 1984, 1987, 1990, 1993, 1996 and these surveys became biennial for the 1999-2003 surveys. The surveys provide much information on Pacific ocean perch, including an abundance index, age composition, and growth characteristics. The surveys are theoretically an estimate of absolute biomass, but I treat them as an index in the stock assessment. The triennial surveys covered all areas of the Gulf of Alaska out to a depth of 500 m (to 1,000 m in some surveys), but the 2001 survey did not sample the eastern Gulf of Alaska. Other, less comprehensive trawl surveys were periodically conducted before 1984 in the Gulf of Alaska, and these have also provided information on age and size composition of slope rockfish. Summaries of biomass estimates from the 2003 trawl survey and comparative estimates from the 1984 to 2003 surveys are provided in Table 1.4.

1.2.2.2 Comparison of Trawl Surveys in 1984-2003

Gulfwide biomass estimates for Pacific ocean perch are shown in Table 1.4. Gulfwide biomass estimates and 95% confidence intervals are also shown graphically in Figure 1.2. The 1984 survey results should be treated with some caution, as a different survey design was used in the eastern Gulf of Alaska. Also, much of the survey effort in 1984 and 1987 was by Japanese vessels that used a very different net design than what has been the standard used by U.S. vessels throughout the surveys. To deal with this problem, fishing power comparisons of rockfish catches have been done for the various vessels used in the surveys (for a discussion see Heifetz et al. 1994). Results of these comparisons have been incorporated into the biomass estimates listed here, and the estimates are believed to be the best available. Even so, the reader should be aware that use of Japanese vessels in 1984 and 1987 does introduce an element of uncertainty as to the standardization of these two surveys.

The biomass estimates for Pacific ocean perch have been extremely variable in recent surveys (Figure 1.2). Such wide fluctuations in biomass do not seem reasonable given the slow growth and low natural mortality rates of POP. Large catches of an aggregated species like Pacific ocean perch in just a few individual hauls can greatly influence biomass estimates and may be a source of much variability. Anomalously large catches have especially affected the biomass estimates for Pacific ocean perch in the 1999 and 2001 surveys. In past SAFE reports, I have also speculated that a change in availability of rockfish to the survey, caused by unknown behavioral or environmental factors, may explain some of the observed variation in biomass. It seems prudent to repeat this speculation in the present report, while acknowledging that until more is known about rockfish behavior, the actual cause of changes in biomass estimates will remain the subject of conjecture. Ongoing research has focused on improving rockfish survey biomass estimates using alternative sampling designs (Quinn et al. 1999, Hanselman et al. 2001, Hanselman et al. 2003). Research on the utility of using hydroacoustics to gain survey precision is also underway.

Biomass estimates of Pacific ocean perch were relatively low in 1984 to 1990, increased markedly in both 1993 and 1996, and became substantially higher in 1999 and 2001 with much uncertainty. Biomass estimates in 2003 have less sampling error with a total similar to the 1993 estimate indicating that the large estimates from 1996-2001 may have been a result of a few anomalous catches. To examine these changes in more detail, the biomass estimates for Pacific ocean perch in each statistical area, along with Gulfwide 95% confidence intervals, are presented in Table 1.4. The large rise in 1993, which the confidence intervals indicate was statistically significant compared with 1990, was primarily the result of big increases in biomass in the Central and Western Gulf of Alaska. The Kodiak area increased greater than tenfold, from 15,221 mt in 1990 to 154,013 mt in 1993. The 1996 survey showed continued biomass increases in all areas, especially Kodiak, which more than doubled compared with 1993. In 1999, there was a substantial decline in biomass in all areas except Chirikof, where a single large catch resulted in a very large biomass estimate. In 2001, the biomass estimates in both the Shumagin and Kodiak areas were the highest of all the surveys. In particular, the biomass in Shumagin was much greater than in previous years; as discussed previously, the increased biomass here can be attributed to very large catches in two hauls. In 2003 the estimated biomass in all areas except for Chirikof decreased, where Chirikof returned from a decade low to an average value.

1.2.2.3 Age Compositions

Ages were determined from the break-and-burn method (Chilton and Beamish 1982). The survey age compositions from 1984-1999 surveys showed that although the fish ranged in age up to 84 years, most of the population was relatively young; mean population age was 11.2 years in 1996 and 13.9 years in 1999 (Table 1.5). The first four surveys identified a relatively strong 1976 year class and also showed a period of very weak year classes prior to 1976 (Figure 1.5). The weak year classes of the early 1970's may have delayed recovery of Pacific ocean perch populations after they were depleted by the foreign fishery. The survey age data from 1990-1999 data suggested that there was a period of large year classes from 1986-1989. In 1990-1993 the 1986 year class looked very strong. Beginning in 1996 and continuing in 1999 survey ages, the 1987 and 1988 year classes became more abundant than the 1986 year class. Rockfish are difficult to age, especially as they grow older, and perhaps some of the fish have been categorized into adjacent age classes between surveys. Alternatively, these year classes were not available to the survey until much later than the 1986 year class. Recruitment of the stronger year classes from the late 1980s probably has accounted for much of the increase in the estimated biomass for Pacific ocean perch in recent surveys.

1.2.2.4 Survey Size Compositions

Gulfwide population size compositions for Pacific ocean perch are shown in Figure 1.6. The size composition for Pacific ocean perch in 2001 was bimodal, which differed from the unimodal compositions in 1993, 1996, and 1999. The 2001 survey showed a large number of relatively small fish, ~32 cm fork length which may indicate recruitment in the early 90's, together with another mode at ~38 cm. Compared to the previous survey years, both 2001 and 2003 showed a much higher proportion of small fish compared to the number of fish in the pooled class of 39+ cm. This could be from good recruitment or from fishing down of larger fish. Survey size data is used in constructing the age-length matrix, but not used in the model fitting phase.

1.3 Analytic Approach

1.3.1 Model Structure

For the third year, I present results for Pacific ocean perch based on an age-structured model using AD Model Builder software (Otter Research Ltd 2000). Previously the stock assessment was based on an age-structured model using stock synthesis (Methot 1990). The assessment model used for Pacific ocean perch is based on a generic rockfish model developed in a workshop held in February 2001[3]. The generic

[3] Rockfish Modeling Workshop, NMFS Auke Bay Laboratory, 11305 Glacier Hwy., Juneau, AK. February, 2001.

rockfish model builds from the northern rockfish model (Courtney et al. 1999). Four changes were made to the northern rockfish model during construction of the generic rockfish model. Fishery age compositions and associated likelihood components were added. The spawner-recruit relationship was removed from the estimation of beginning biomass (B_0). Survey catchability, q, was computed relative to survey selectivity standardized to a maximum of one (full selectivity), rather than to survey selectivity standardized to an average of one (average selectivity). The penalties for deviations from reasonable fishing mortality parameter estimates were modified. These fishing mortality deviation and regularity penalties are part of the internal model structure and are designed to speed up model convergence. The result was a separable age-structured model with allowance for size composition data that was adaptable to several rockfish species. The parameters, population dynamics and equations of the model are described in Box 1.1. Since its initial adaptation in 2001, the model's attributes have been explored and several new changes are proposed below.

1.3.2 Parameters Estimated Independently

The estimate of natural mortality (M) was based on catch curve analysis to determine Z. Estimates of Z could be considered as an upper bound for M. Estimates of Z for Pacific ocean perch from Archibald et al. (1981) were from populations considered to be lightly exploited and thus are considered reasonable estimates of M, yielding a value of ~0.05. In some model scenarios I estimate M, but use 0.05 as the mean of a prior distribution.

Recently, new information on female age and size at 50% maturity has become available for Pacific ocean perch from a study in the Gulf of Alaska that is based on the currently accepted break-and-burn method of determining age from otoliths (Lunsford 2000). These data are summarized below (size is in cm fork length and age is in years), and the full maturity schedule is in Table 1.6:

Sample size	Size at 50% maturity	Age at 50% maturity
802	35.7	10

A von Bertalanffy growth curve was fitted to survey length at age data from 1984-1999. Sexes were combined. A length at age transition matrix was then constructed by adding normal error with a standard deviation equal to the survey data for the probability of different ages for each size class. Two new matrices were constructed for the two alternative models considered in this year's SAFE. A second matrix was constructed to represent a lower growth rate in the 1960s. The estimated parameters for the growth curve are shown below:

$$L_\infty = 41.4 \text{ cm} \quad \kappa = 0.19 \quad t_0 = -0.47 \quad n = 9336$$

where L_∞ is the average maximum length, κ is the shape parameter of the curve, t_0 is the intercept, and n is the sample size.

Weight-at-age was constructed with weight at age data from the same data set as the length at age. The estimated growth parameters are shown below. A correction of $(W_\infty - W_a)/2$ was used for the weight of the pooled ages (Schnute et al. 2001).

$$W_\infty = 984 \text{ g} \quad a = 0.0004 \quad b = 2.45 \quad n = 3592$$

where W_∞ is maximum average weight, a is the linear growth parameter, b is the allometric growth parameter, and n is the sample size.

Aging error matrices were constructed by assuming that the break-and-burn ages were unbiased but had an increasing normal error as age increased.

1.3.3 Parameters estimated conditionally

Parameters estimated conditionally include but are not limited to: catchability, selectivity (up to full selectivity) for survey and fishery, recruitment deviations, mean recruitment, fishing mortality, and spawners per recruit levels. Other parameters are described in Box 1.1.

1.3.4 Uncertainty

Evaluation of model uncertainty has recently become an integral part of the "precautionary approach" in fisheries management. In complex stock assessment models such as this model, evaluating the level of uncertainty is difficult. One way is to examine the standard errors of parameter estimates from the Maximum Likelihood (ML) approach derived from the Hessian matrix. While these standard errors give some measure of variability of individual parameters, they often underestimate their variance and assume that the joint distribution is multivariate normal. An alternative approach is to examine parameter distributions through Markov Chain Monte Carlo (MCMC) methods (Gelman et al. 1995). When treated this way, the stock assessment is a large Bayesian model, which includes informative (e.g., lognormal natural mortality with a small CV) and noninformative (or nearly so, such as a parameter bounded between 0 and 10) prior distributions. In the models presented in this SAFE report, the number of parameters estimated was between 131 and 134. In a low-dimensional model, an analytical solution might be possible, but in one with this many parameters, an analytical solution is intractable. Therefore, I use MCMC methods to estimate the Bayesian posterior distribution for these parameters. The basic premise is to use a Markov chain to simulate a random walk through the parameter space which will eventually converge to a stationary distribution which approximates the posterior distribution. Determining whether a particular chain has converged to this stationary distribution can be complicated, but generally if allowed to run long enough, it will converge. The "burn-in" is a set of iterations removed at the beginning of the chain. In these simulations, I removed the first 500,000 iterations out of 5,000,000 and "thinned" the chain to one value out of every thousand, leaving a sample distribution of 4,500. Further assurance that the chain had converged was to compare the mean of the first half of the chain with the second half after removing the "burn-in" and "thinning." Because these two values were similar, I concluded that convergence had been attained. I use these MCMC methods to provide further evaluation of uncertainty in the results below and to show examples of key parameter posterior distributions (Figures 1.7, 1.8).

Parameter definitions	**Box 1.1. AD Model Builder POP Model Description**
y	Year
a	Age classes
l	Length classes
w_a	Vector of estimated weight at age, $a_0 \rightarrow a_+$
m_a	Vector of estimated maturity at age, $a_0 \rightarrow a_+$
a_0	Age at first recruitment
a_+	Age when age classes are pooled
μ_r	Average annual recruitment, log-scale estimation
μ_f	Average fishing mortality
ϕ_y	Annual fishing mortality deviation
τ_y	Annual recruitment deviation
σ_r	Recruitment standard deviation
fs_a	Vector of selectivities at age for fishery, $a_0 \rightarrow a_+$
ss_a	Vector of selectivities at age for survey, $a_0 \rightarrow a_+$
M	Natural mortality, log-scale estimation
$F_{y,a}$	Fishing mortality for year y and age class a $(fs_a \mu_f e^{\varepsilon})$
$Z_{y,a}$	Total mortality for year y and age class a $(=F_{y,a}+M)$
$\varepsilon_{y,a}$	Residuals from year to year mortality fluctuations
$T_{a,a'}$	Aging error matrix
$T_{a,l}$	Age to length transition matrix
q	Survey catchability coefficient
SB_y	Spawning biomass in year y, $(=m_a\, w_a\, N_{y,a})$
M_{prior}	Prior mean for natural mortality
q_{prior}	Prior mean for catchability coefficient
$\sigma_{r(prior)}$	Prior mean for recruitment variance
σ^2_M	Prior CV for natural mortality
σ^2_q	Prior CV for catchability coefficient
$\sigma^2_{\sigma_r}$	Prior CV for recruitment deviations

Box 1.1 (Continued)

Equations describing the observed data

$$\hat{C}_y = \sum_a \frac{N_{y,a} * F_{y,a} * \left(1 - e^{-Z_{y,a}}\right)}{Z_{y,a}} * w_a$$

Catch equation

$$\hat{I}_y = q \; * \sum_a N_{y,a} * \frac{ss_a}{\max\left(ss_a\right)} * w_a$$

Survey biomass index (mt)

$$\hat{P}_{y,a'} = \sum_a \left(\frac{N_{y,a} * ss_a}{\sum_a N_{y,a} * ss_a} \right) * T_{a,a'}$$

Survey age distribution

Proportion at age

$$\hat{P}_{y,l} = \sum_a \left(\frac{N_{y,a} * ss_a}{\sum_a N_{y,a} * ss_a} \right) * T_{a,l}$$

Survey length distribution

Proportion at length

$$\hat{P}_{y,a'} = \sum_a \left(\frac{\hat{C}_{y,a}}{\sum_a \hat{C}_{y,a}} \right) * T_{a,a'}$$

Fishery age composition

Proportion at age

$$\hat{P}_{y,l} = \sum_a \left(\frac{\hat{C}_{y,a}}{\sum_a \hat{C}_{y,a}} \right) * T_{a,l}$$

Fishery length composition

Proportion at length

Equations describing population dynamics

Start year

$$N_a = \begin{cases} e^{\left(\mu_r + \tau_{styr - a_0 - a - 1}\right)}, & a = a_0 \\ e^{\left(\mu_r + \tau_{styr - a_0 - a - 1}\right)} e^{-(a - a_0)M}, & a_0 < a < a_+ \\ \dfrac{e^{\left(\mu_r\right)} e^{-(a - a_0)M}}{\left(1 - e^{-M}\right)}, & a = a_+ \end{cases}$$

Number at age of recruitment

Number at ages between recruitment and pooled age class

Number in pooled age class

Subsequent years

$$N_{y,a} = \begin{cases} e^{\left(\mu_r + \tau_y\right)}, & a = a_0 \\ N_{y-1,a-1} * e^{-Z_{y-1,a-1}}, & a_0 < a < a_+ \\ N_{y-1,a-1} * e^{-Z_{y-1,a-1}} + N_{y-1,a} * e^{-Z_{y-1,a}}, & a = a_+ \end{cases}$$

Number at age of recruitment

Number at ages between recruitment and pooled age class

Number in pooled age class

Formulae for likelihood components	**Box 1.1 (Continued)**
$L_1 = \lambda_1 \sum_y \left(\ln \left[\frac{C_y + 0.01}{\hat{C}_y + 0.01} \right] \right)^2$	Catch likelihood
$L_2 = \lambda_2 \sum_y \frac{\left(I_y - \hat{I}_y\right)^2}{2 * \hat{\sigma}^2 \left(I_y\right)}$	Survey biomass index likelihood
$L_3 = \lambda_3 \sum_{styr}^{endyr} -n^*{}_y \sum_a^{a+} \left(P_{y,a} + 0.001\right) * \ln\left(\hat{P}_{y,a} + 0.001\right)$	Fishery age composition likelihood ($n^*{}_y$ =sample size, standardized to maximum of 100)
$L_4 = \lambda_4 \sum_{styr}^{endyr} -n^*{}_y \sum_l^{l+} \left(P_{y,l} + 0.001\right) * \ln\left(\hat{P}_{y,l} + 0.001\right)$	Fishery length composition likelihood
$L_5 = \lambda_5 \sum_{styr}^{endyr} -n^*{}_y \sum_a^{a+} \left(P_{y,a} + 0.001\right) * \ln\left(\hat{P}_{y,a} + 0.001\right)$	Survey age composition likelihood
$L_6 = \lambda_6 \sum_{styr}^{endyr} -n^*{}_y \sum_l^{l+} \left(P_{y,l} + 0.001\right) * \ln\left(\hat{P}_{y,l} + 0.001\right)$	Survey size composition likelihood
$L_7 = \frac{1}{2\sigma_M^2} \left(M / M_{prior} \right)^2$	Penalty on deviation from prior distribution of natural mortality
$L_8 = \frac{1}{2\sigma_q^2} \left(q / q_{prior} \right)^2$	Penalty on deviation from prior distribution of catchability coefficient
$L_9 = \frac{1}{2\sigma_{\sigma_r}^2} \left(\sigma_r / \sigma_{r(prior)} \right)^2$	Penalty on deviation from prior distribution of recruitment deviations
$L_{10} = \lambda_{10} \left[\frac{1}{2 * \sigma_r^2} \sum_y \tau_y^2 + n_y * \ln\left(\sigma_r\right) \right]$	Penalty on recruitment deviations
$L_{11} = \lambda_{11} \sum_y \varepsilon_y^2$	Fishing mortality regularity penalty
$L_{12} = \lambda_{12} \left[\ln(\overline{ss_a})^2 + \ln(\overline{fs_a})^2 \right]$	Average selectivity penalty (attempts to keep average selectivity near 1)
$L_{13} = \lambda_{13} \left[\sum_{a_0}^{a_+} \left(ss_a - ss_{a+1}\right)^2 + \sum_{a_0}^{a_+} \left(fs_a - fs_{a+1}\right)^2 \right]$	Selectivity dome-shapedness penalty – only penalizes when the next age's selectivity is lower than the previous (penalizes a downward selectivity curve at older ages)
$L_{14} = \lambda_{14} \left[\sum_{a_0}^{a_+} \left(SD(ss_a - ss_{a+1})\right)^2 + \sum_{a_0}^{a_+} \left(SD(fs_a - fs_{a+1})\right)^2 \right]$	Selectivity regularity penalty (penalizes large deviations from adjacent selectivities by summing the square of second differences
$L_{total} = \sum_{i=1}^{14} L_i$	Total objective function value

1.4 Model Evaluation

1.4.1 Alternative Models

1.4.1.1 Base Model

This model was the base model that has been used in the previous two slope rockfish SAFE documents for Pacific ocean perch. Except for catch data, the base model was run with all data components given a likelihood weight of one and both survey and fishery selectivity patterns constrained to be approximately asymptotic. The catch likelihood was given a weight of 50 in all model runs. Each year of data components was weighted within a likelihood component by computing the square root of the sample size and scaling it to a maximum of 100. Table 1.7 summarizes the results from the base model and the new alternative models. Figures 1.9 to 1.14 show some of the results for the base model. For this base model the fit to survey biomass was poor for the more recent surveys. In addition the fits to some of the survey age compositions were not very good. The predicted fits to fishery length compositions are poor (Figure 1.11) and have a large influence on overall model fits. This was partly because the length compositions are the longest time series in the model. I surmise that this poor fit was also due to an inaccurate length at age transition matrix. The base model also relies heavily on penalties that caused peculiar distributions in the Markov Chain Monte Carlo (MCMC) outputs explored in last year's SAFE (Heifetz et al. 2002). An example of this is in Figure 1.10 where the predicted total biomass from the model is outside of the 95% MCMC confidence interval. Further discussion of MCMC methods used for assessing uncertainty was presented in Section 1.3.4. The next model also explores lowering or removing these constraints.

1.4.1.2 Model 2

In model two I made extensive changes to the base model. The large likelihood component of the length frequency data in the base model led to further examination of the current length at age matrix. This revealed some unlikely components of the base model matrix. Primarily, the matrix predicts that an older fish would fall into an unrealistically small size class. This matrix was based on limited age data from when the stock synthesis approach was used. In model two, a new length at age matrix was constructed using a slightly different method than the previous SAFEs that alleviates this unreasonable probability distribution. A new LVB model was fit to the data using survey data from 1984-1999. The matrix was then constructed using the predicted lengths at age and observed standard deviation at age. The new matrix lowers the effect of the size data on the objective function (Table 1.7) and provides much better fits to the data. I remove one year of size data (1978) which has an unusual distribution (Table 1.2) and exerts leverage on the model even though it has a small sample size.

I estimate natural mortality (M) but use an informative prior (lognormal, mean=0.05, σ=0.01) which admits a little uncertainty, but constrains it from extreme values. Figure 1.12 of the base model predicts that fishing mortality was too low in the past, considering that 1.7 million mts were removed between 1963 and 1978. The fully-selected F of 0.4 predicted in the base model for 1965 with a 350,000 mt catch translates to 1.1 million mts, while the base model predicts that there were only ~800,000 mts of exploitable biomass. Hence, I lower the fishing mortality regularity penalty from 1 to 0.1 which was consistent with other AD Model Builder assessments (e.g., sablefish and BSAI Pacific ocean perch). Figure 1.13 shows that estimated recruitments over time have been reasonably consistent according to the base model. This regularity was unexpected for rockfish considering that it is commonly believed that their populations are characterized by rare large recruitment levels. The prior mean for the recruitment deviation parameter (σ_r) in the current model was 0.9. This value, which implies a CV for log-recruitment of 25%, seems low considering the current theory of sporadic recruitment. Additionally, Figure 1.7 shows the MCMC distribution of the recruitment deviation parameter and shows the previous bound set on it was unreasonable, with much of the mass truncated at two. Therefore, I set the mean of the σ_r prior distribution to be 1.7 with a CV of 0.2 in this model (1.7 is roughly the mode of the MCMC distributions in Figure 1.8 and other rockfish species) and increase the upper boundary on σ_r to ten from two. A correction factor was added to the weight at age relationship to compensate for the pooling of ages after age 25 using a method suggested by Schnute et al. (2001). Other penalties in the model were lowered to one from the base model as can be seen in Table 1.7.

1.4.1.3 Models 3-5

Models 3-5 add an additional length at age matrix. The biomass in the 1960s was likely much larger than present. Since POP seem to inhabit small optimum areas, I suspect this substantial reduction in the population caused a concurrent density dependent increase in growth. Evidence of this suspicion can be seen when examining the fishery size data. In the size data from 1963-1977, the weighted-average size was 34 cm while the second set of size data from 1990-1999 has a weighted-average of 36.5 cm, representing approximately a ~6% increase in average growth. Since the length at age matrix applied to these data was based only on recent length at age data, this matrix will give poor results when applied to the older size data. I constructed a slower-growth length at age matrix to use for the size data from 1963-1977 that reflects that older fish have a smaller size. The method here was simple: decrease the length-at-age by six percent, then refit the LVB model and use the resulting matrix for predicting those years. This resulted in a better fit to the fishery size data, survey age data and a better overall fit of the model. For comparison Model 3 shows a fixed natural mortality at 0.05, Model 4 shows a fixed M $(M=0.05)$ and q constrained to one. Model 5 is the "full" model that estimates q and M simultaneously.

1.4.2 Model Comparison

I compare stock assessment results for the five different model configurations above:

Model 1 - Base model from 2002 SAFE

Model 2 - New length at age transition matrix applied, penalties reduced, new weight-at-age

Model 3 – 2nd length at age transition matrix applied to fishery lengths 1963-77, M fixed at 0.05

Model 4 – Model 3 with M fixed at 0.05 and q constrained to equal 1.

Model 5 –Model 3 with q and M both estimated.

Models 2-5 all have significantly better fits than the base model. The changes made in Model 2 make it a reasonable choice, but does not fit the data as well as Models 3-5. The objective function was reduced significantly for this model and the results in general are more appealing than the base model but some results are unexpected (e.g. a recruitment of ~1 billion fish in the first year of the model and an equivalent spawning biomass at the beginning of the time series as at the end.)

Model 4 produced a better fit to the data than the base model and model 2, but was less stable, requiring the fishing mortality regularity penalty to be raised to 0.2 from 0.1 for convergence. Models 3 and 5 have the best overall fit, with 5 fitting slightly better and providing more reasonable estimates of q and B_{2004} than Model 3. Even though the penalties for selectivity smoothness were lowered, selectivities in model 5 were still reasonable (Table 1.6). Models 3 and 5 produced reasonable estimates after lowering all the penalties to quantities that have little effect on the model, indicating increased stability. Overall, model 5 has the best properties of the alternatives and I recommend model 5 for setting the ABC in 2004.

1.5 Model Results

Model 5 shows a much improved fit to age and length data (Figures 1.3 to 1.6). An example of the improved fit was provided by comparing the length predictions in Figure 1.4 with the base model predictions in Figure 1.11. MCMC confidence intervals around predicted biomass (Figure 1.15 and 1.16) show a more realistic reflection of uncertainty around recent biomass predictions than the base model (Figure 1.9 and 1.10). There are very tight confidence intervals around recent estimates from the base model, with the model estimate outside of the confidence intervals. This model, when compared to the base model has a 55% smaller objective function value and 63% smaller portion of the objective function attributed to the data fits. Table 1.7 shows a summary of the main results for model 5. Additional results for model 5 are shown in Figures 1.15 to 1.21; fits to the data are shown in Figures 1.2 to 1.6.

Model 5 suggests that there was a group of stronger recruitments in the late 1980s, peaking with a very large recruitment (age-2) in 1989. Before then, the model suggests there were no other major recruitment events since the 1960s. MCMC confidence intervals around recruitments reflect much uncertainty around these estimated recruitments, particularly in recent years (Figure 1.20). Marginal posterior distributions from the MCMC integration suggest that the estimates could be quite different from the mode and that prior distributions did not particularly affect the estimates except for M (Figure 1.8). The tight posterior distribution of natural mortality was due to its prior CV of 0.01 which was necessary to prevent very large estimates of M, which in turn would produce low estimates of q.

I suggest that in the face of uncertainty, it is preferable to be more conservative and accept a moderately high estimate of q rather than move to a much higher estimate of natural mortality. In a model with this many parameters q cannot be considered as a true measure of trawl catchability, but as a scaling factor that is affected by other data in the model. One possibility is that in the years the trawl survey has one or two tows that are an order of magnitude larger than the rest of the tows, these tows are translated into unexpected jumps in certain age or length classes. This would lead to q rising to compensate for this increase in catchability. In model 5, if the natural mortality was allowed to rise to 0.075, this equates to a q of about 1.

From the MCMC chains described in Section 1.5.3, I summarized the posterior densities of key parameters for the recommended model using histograms (Figure 1.8) and confidence regions (Table 1.8). I also used these posterior distributions to show uncertainty around time series estimates such as total biomass, spawning biomass and recruitment (Figs. 1.9, 1.10, 1.15, 1.16).

Table 1.8 shows the maximum likelihood estimate (MLE) of key parameters with their corresponding MLE standard deviation derived from the Hessian matrix. Also shown are the MCMC standard deviation and the corresponding Bayesian 95%

confidence intervals (BCI). The MLE and MCMC standard deviations are similar for q, M and F_{40}, but the MCMC standard deviations are much larger for the estimates of B_{2004}, ABC and σ_r (recruitment deviation). These larger standard deviations indicate that these parameters are more uncertain than indicated by the standard modeling, especially in the case of σ_r in which the MLE estimate was far out of the Bayesian confidence intervals. This highlights a concern that σ_r requires a fairly informative prior distribution since it is confounded with available data on recruitment variability. To illustrate this problem, imagine a stock that truly has variable recruitment. If this stock lacks age data (or the data are very noisy), then the modal estimate of σ_r is near zero. The distribution of ABC and spawning biomass are highly skewed, indicating possibilities of much higher biomass estimates (also see Figure 1.8).

I selected the results from Model 5, a new model, as the basis for my recommendations for ABC and overfishing. The ABC for this year's assessment is similar to last year's assessment using $F_{40\%}$. Recently, the use of $F_{40\%}$ has come into question for rockfish in a NPFMC harvest strategy review (Goodman et al. 2002). Adoption of a more conservative harvest strategy such as $F_{50\%}$ has been suggested for West Coast rockfish in recent literature (Dorn 2002, Ianelli 2002, Hilborn et al. 2002). I do not feel these papers apply particularly well to Gulf of Alaska rockfish, which likely are healthier and more productive than West Coast stocks (Dorn 2002). Therefore I recommend continuing to harvest at $F_{40\%}$ unless new information suggests otherwise.

1.6 Harvest Alternatives

1.6.1 Harvest Alternatives

Several alternative model configurations were evaluated in section 1.7.1. ABCs from these alternative models ranged from 9,400 – 19,877 mt. I recommend that the ABC from model 5 be used for the 2004 fishery. The management path from Model 5 in Figure 1.20 suggests that management is on track and moving the stock into the 'optimum' quadrant where $B_{now}/B_{40\%}$ has recently exceeded one again for the first time since the 1960s. $F_{now}/F_{40\%}$ continues to stay below one. Based on model 5, the spawning biomass in 2004, B_{2004}, is 95,760 mt. $B_{40\%}$ is 89,699 mt which was determined from average recruitment of the 1977-97 year-classes (Table 1.9). Since B_{2004} is greater than $B_{40\%}$, the computation in tier 3a [i.e., $F_{ABC} = F_{40\%}$] is used to determine the maximum value of F_{ABC} resulting in an ABC of 13,340 mt. I expected to recommend a larger ABC this year before receiving the 2003 survey biomass estimate. Using last survey's biomass estimate as a placeholder, the recommended model was predicting a much higher ABC (18,112). This year's survey biomass estimate came in much lower and more precise than recent years, resulting in a return to approximately the base model's ABC from last year. I recommend that the ABC for Pacific ocean perch for 2004 fishery in the Gulf of Alaska be set at 13,340 mt.

1.6.2 Area Allocation of Harvests

Prior to the 1996 fishery, the apportionment of ABC among areas was determined from distribution of biomass based on the average proportion of exploitable biomass by area in the most recent three triennial trawl surveys. For the 1996 fishery, an alternative method of apportionment was recommended by the Plan Team and accepted by the Council. Recognizing the uncertainty in estimation of biomass yet wanting to adapt to current information, the Plan Team chose to employ a method of weighting prior surveys based on the relative proportion of variability attributed to survey error. Assuming that survey error contributes 2/3 of the total variability in predicting the distribution of biomass (a reasonable assumption), the weight of a prior survey should be 2/3 the weight of the preceding survey. This results in weights of 4:6:9 for the 1999, 2001, and 2003 surveys, respectively and apportionments of 19% for the Western area, 63 % for the Central area, and 18% for the Eastern area (Table 1.10). This results in recommended ABC's of 2,520 mt for the Western area, 8,390 mt for the Central area, and 2,430 mt for the Eastern area.

Amendment 41 prohibited trawling in the Eastern area east of 140° W longitude. In the past, the Plan Team has calculated an apportionment for the West Yakutat area that is still open to trawling (between 147°W and 140°W). I calculated this apportionment using the ratio of estimated biomass in the closed area and open area. This calculation was based on the team's previous recommendation that I use the weighted average of the upper 95% confidence interval for the W. Yakutat. I computed this interval this year using the weighted average of the ratio for 1996, 1999 and 2003. I calculated the upper 95% confidence interval using the variance of the 1996-2003 ratios for the weighted variance estimate. This resulted in a similar ratio as last year of 0.34. This results in an apportionment to the W. Yakutat area of 830 mt which would leave 1600 mt unharvested in the Eastern Gulf.

1.6.3 Overfishing Definition

Based on the definitions for overfishing in Amendment 44 in tier 3a (i.e., $F_{OFL} = F_{35\%}=0.071$), overfishing was set equal to 15,840 mt for Pacific ocean perch. The overfishing level was apportioned by area for Pacific ocean perch. Using the apportionment in Section 1.8.3, results in overfishing levels by area of 3,000 mt in the Western area, 9,960 mt in the Central area, and 2,880 mt in the Eastern area.

1.7 Ecosystem Considerations

In general, a determination of ecosystem considerations for slope rockfish was hampered by the lack of biological and habitat information.

1.7.1 Ecosystem Effects on the Stock

Prey availability/abundance trends: similar to many other rockfish species, stock condition of Pacific ocean perch appears to be influenced by periodic abundant year classes. Availability of suitable zooplankton prey items in sufficient quantity for larval or post-larval Pacific ocean perch may be an important determining factor of year class strength. Unfortunately, there is no information on the food habits of larval or post-larval rockfish to help determine possible relationships between prey availability and year class strength; moreover, identification to the species level for field collected larval slope rockfish is difficult. Visual identification is not possible though genetic techniques allow identification to species level for larval slope rockfish (Gharrett et. al 2001). Some juvenile rockfish found in inshore habitat feed on shrimp, amphipods, and other crustaceans, as well as some molluscs and fish (Byerly 2001). Adult Pacific ocean perch feed primarily on euphausiids. Little if anything is known about abundance trends of likely rockfish prey items. Euphausiids are also a major item in the diet of walleye pollock. Changes in the abundance of walleye pollock could lead to a corollary change in the availability of euphausiids, which would then have an impact on Pacific ocean perch.

Predator population trends: Pacific ocean perch are preyed on by a variety of other fish at all life stages, and to some extent marine mammals during late juvenile and adult stages. Whether the impact of any particular predator is significant or dominant is unknown. Predator effects would likely be more important on larval, post-larval, and small juvenile slope rockfish, but information on these life stages and their predators is scarce.

Changes in physical environment: Stronger year classes corresponding to the period around 1977 have been reported for many species of groundfish in the Gulf of Alaska, including Pacific ocean perch, northern rockfish, sablefish, and Pacific cod. Therefore, it appears that environmental conditions may have changed during this period in such a way that survival of young-of-the-year fish increased for many groundfish species, including slope rockfish. Pacific ocean perch appeared to have a strong 1987-88 year classes, and these may be other years when environmental conditions were especially favorable for rockfish species. The environmental mechanism for this increased survival remains unknown. Changes in water temperature and currents could have effect on prey item abundance and success of transition of rockfish from pelagic to demersal stage. Rockfish in early juvenile stages have been found in floating kelp patches which would be subject to ocean currents. Changes in bottom habitat due to natural or anthropogenic causes could alter survival rates by altering available shelter, prey, or other functions.

1.7.2 Fishery Effects on the Ecosystem

Fishery-specific contribution to bycatch of HAPC biota: In the Gulf of Alaska, bottom trawl fisheries for pollock, deepwater flatfish, and Pacific ocean perch account for most of the observed bycatch of coral, while rockfish fisheries account for little of the bycatch of sea anemones or of sea whips and sea pens. The bottom trawl fisheries for Pacific ocean perch and Pacific cod and the pot fishery for Pacific cod accounts for most of the observed bycatch of sponges (Table 1.11).

Fishery-specific concentration of target catch in space and time relative to predator needs in space and time (if known) and relative to spawning components: The directed slope rockfish trawl fisheries begin in July concentrated in known areas of abundance and typically lasts only a few weeks. The recent annual exploitation rates on rockfish are thought to be quite low. Insemination is likely in the fall or winter, and parturition is likely mostly in the spring. Hence, reproductive activities are probably not directly affected by the commercial fishery.

Fishery-specific effects on amount of large size target fish: There is no evidence for targeting large fish since the size-at-age has increased since the beginning of the fishery.

Fishery contribution to discards and offal production: Fishery discard rates for the whole rockfish trawl fishery have declined from 35% in 1997 to 19% in 2002. Arrowtooth flounder comprised 22-46% of these discards.

Fishery-specific effects on age-at-maturity and fecundity of the target fishery: Speculatively, I would expect that if the size-at-age is getting larger, than fecundity is rising and age-at-maturity is decreasing. However, no studies have been conducted to provide evidence of this.

Fishery-specific effects on EFH non-living substrate: Effects on non-living substrate are unknown, but the heavy-duty "rockhopper" trawl gear commonly used in the fishery is suspected to move around rocks and boulders on the bottom.

1.7.3 Data Gaps and Research Priorities

There is little information on larval, post-larval, or early stages slope rockfish. Habitat requirements for larval, post-larval, and early stages are mostly unknown. Habitat requirements for later stage juvenile and adult fish are anecdotal or conjectural. Research needs to be done on the bottom habitat of the major fishing grounds, on what HAPC biota are found on these grounds, and on what impact bottom trawling has on these biota. Additionally, Pacific ocean perch are undersampled by the current survey design. The stock assessment would benefit from additional survey effort and age-reading.

1.8 Summary

A summary of biomass levels, exploitation rates and ABCs for slope Pacific ocean perch is in the following table:

	Model				
	1 Base	2 New size-age matrix, low penalties	3 Model 2 with M fixed @ 0.05 and two size-age matrices	4 Model 3 with q constrained to = 1	5* Full model, estimating M and q
Tier	3a				
Total Biomass (Age 2+)	360,650	384,060	250,510	508,230	285,070
B_{2004} (mt)	120,090	138,385	95,567	166,100	95,765
$B_{0\%}$ (mt)	280,254	290,955	238,918	366406	224,248
$B_{40\%}$ (mt)	112,102	116,382	85,840	146,562	89,699
$B_{35\%}$ (mt)	98,089	101,834	83,622	128,242	78,486
M	0.05	0.06	0.05	0.05	0.06
$F_{40\%}$	0.05	0.06	0.05	0.05	0.06
F_{ABC} (maximum allowable)	0.05	0.06	0.05	0.05	0.06
ABC (mt; maximum allowable)	14,761	18,519	9,406	19,877	13,340

* Recommended for ABC calculation

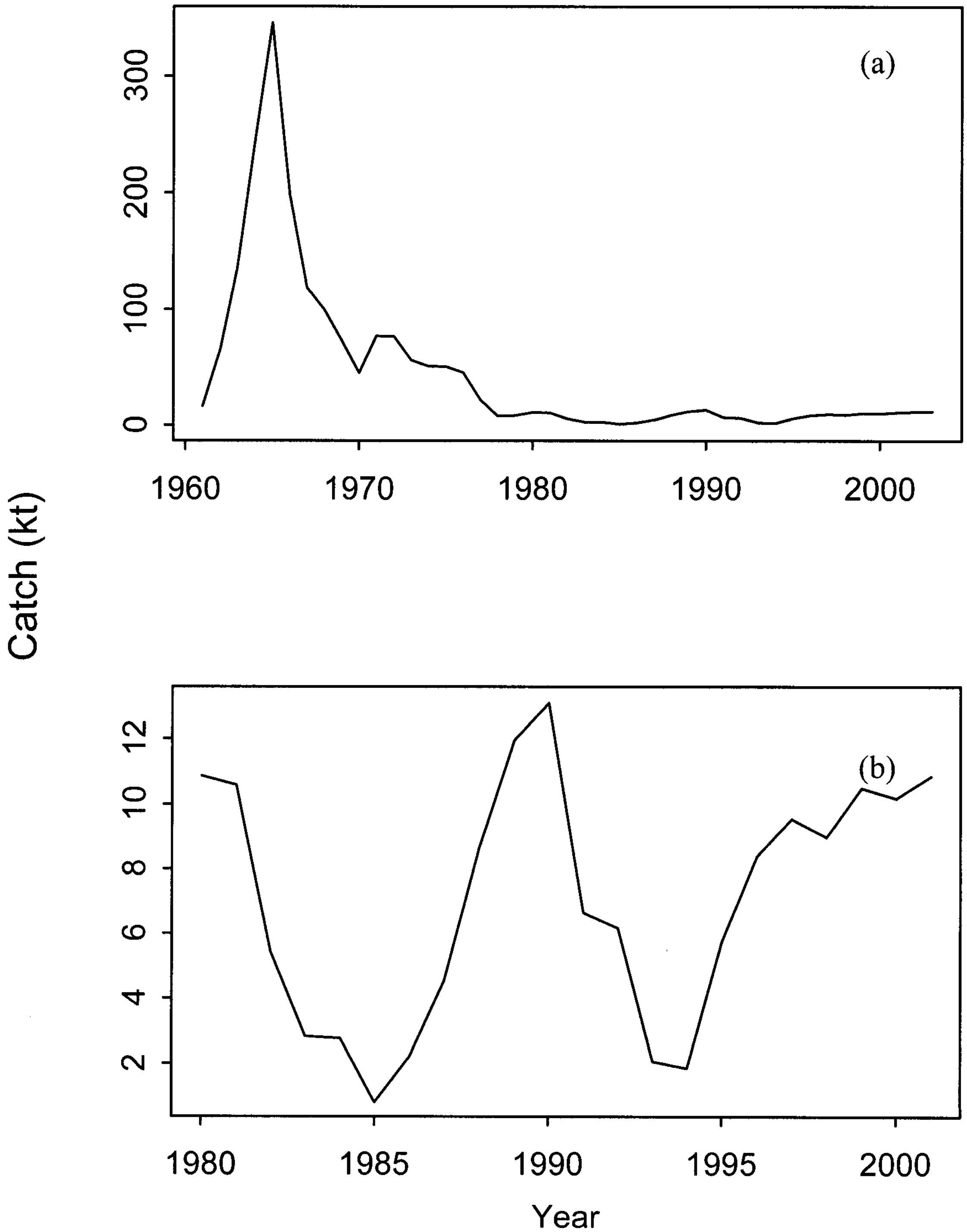


Figure 1.1 Long (a) and short (b) term commercial catches for Gulf of Alaska Pacific ocean perch.

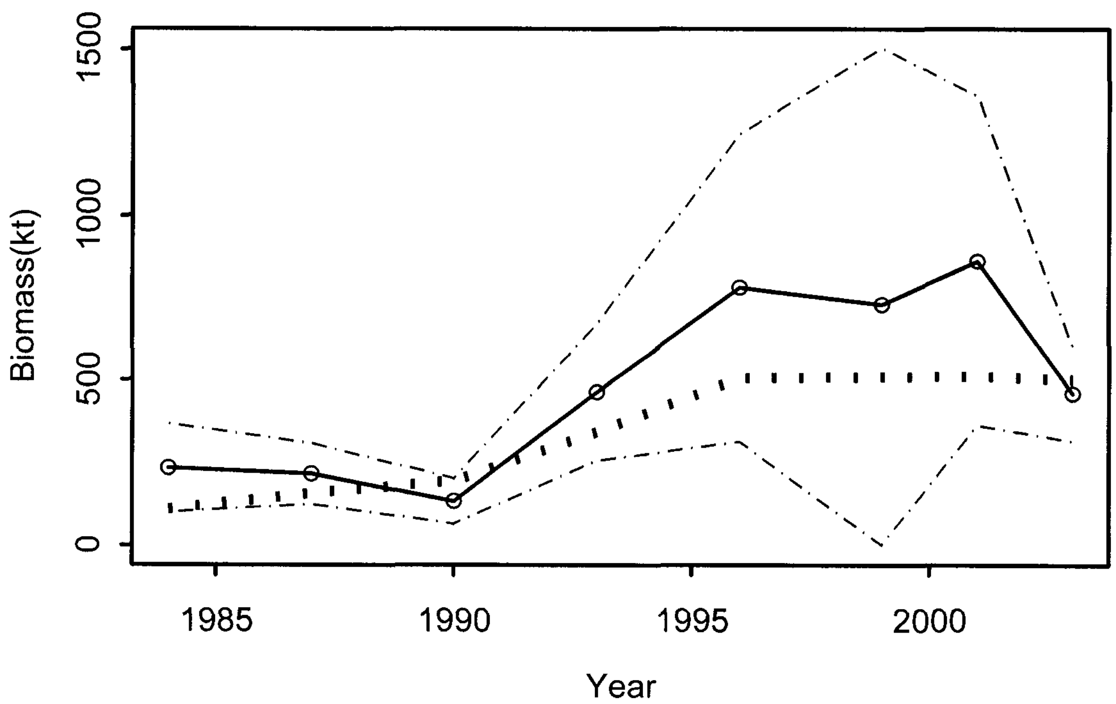


Figure 1.2. Observed and predicted GOA POP survey biomass. Observed biomass=solid line and model predicted biomass=dotted line. Outer dashed lines represent 95% CIs of sampling error for observed biomass.

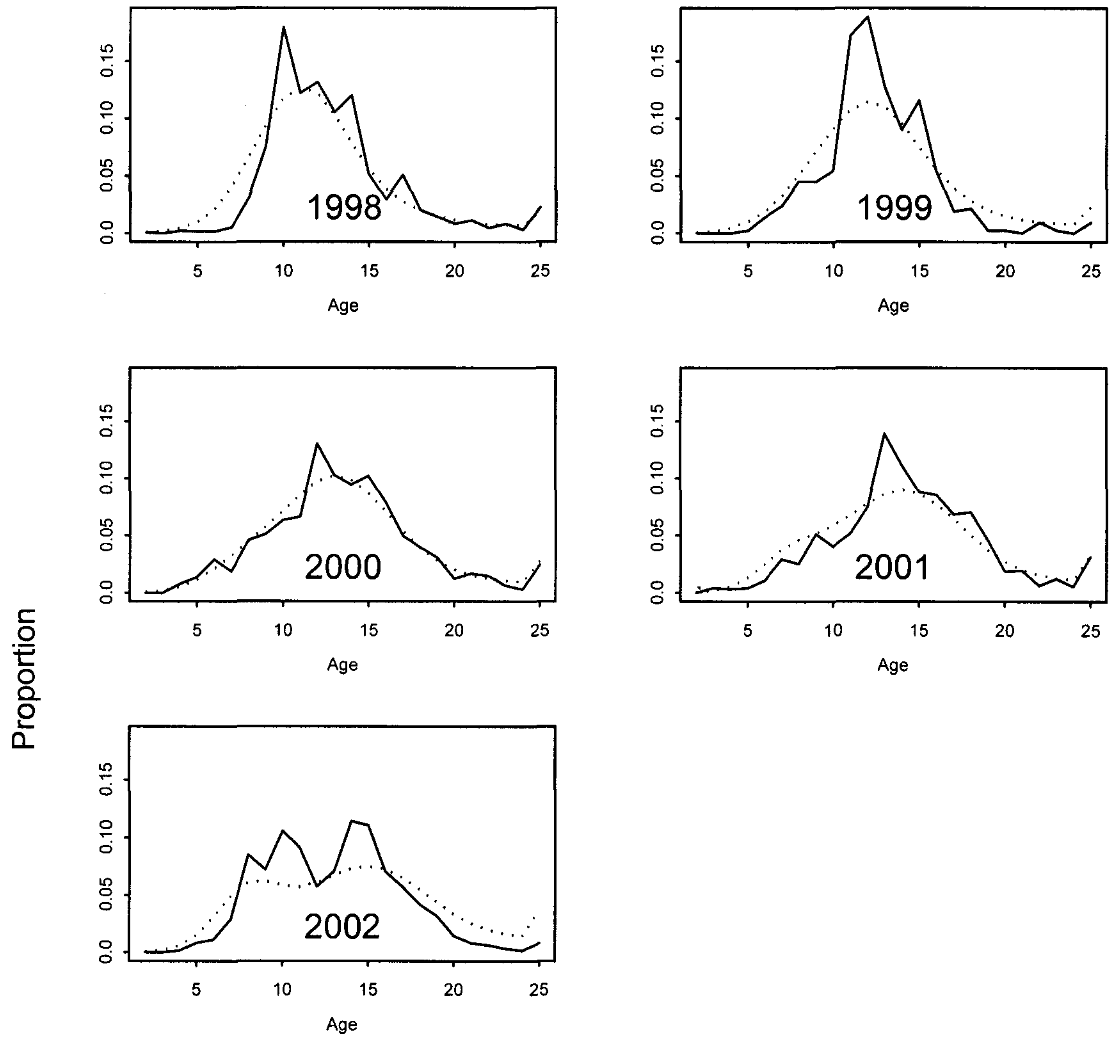


Figure 1.3. Fishery age compositions for GOA Pacific ocean perch. Solid line=observed, Dotted line=model predicted.

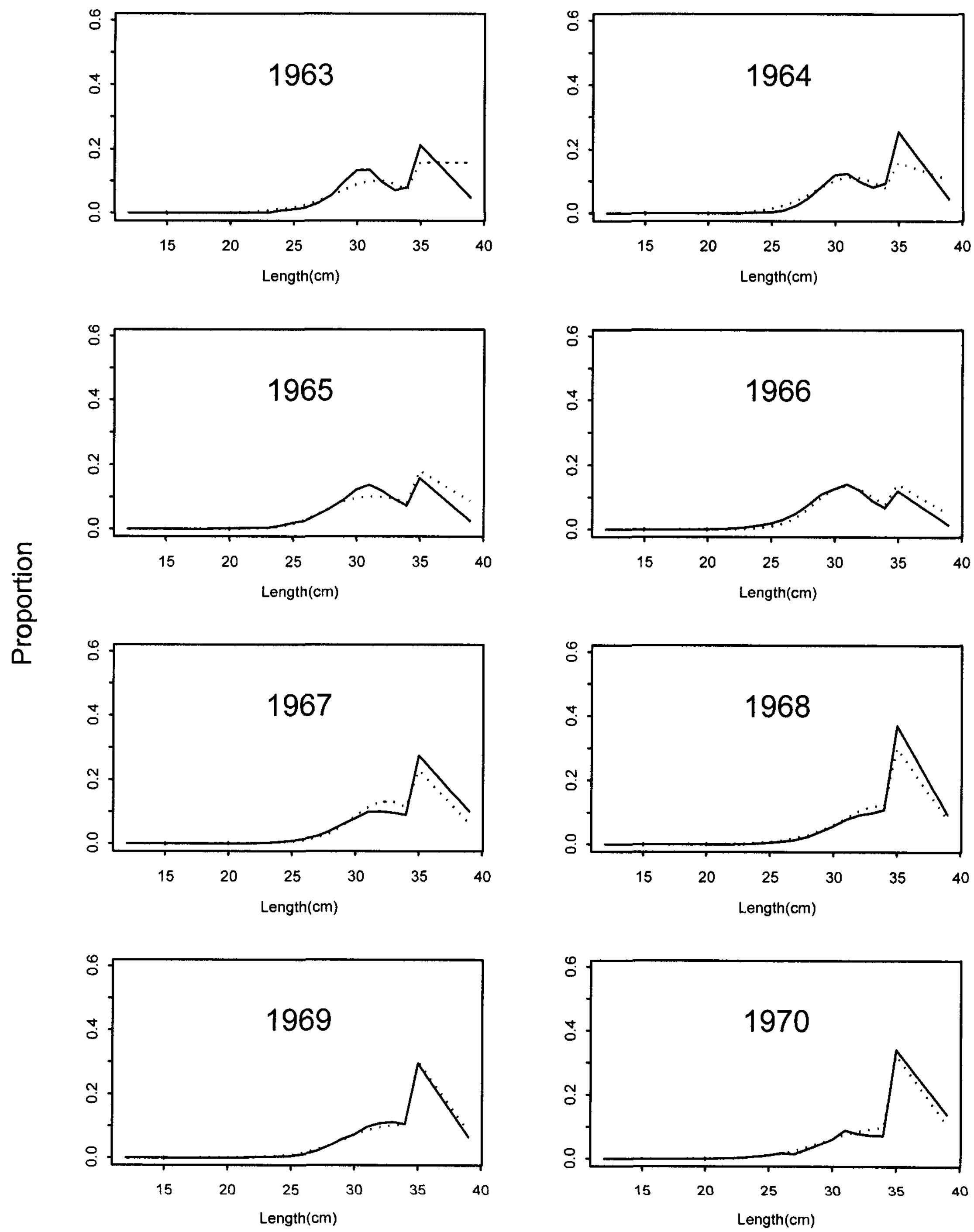


Figure 1.4. Fishery length compositions for GOA POP. Solid line=observed, dotted line=predicted.

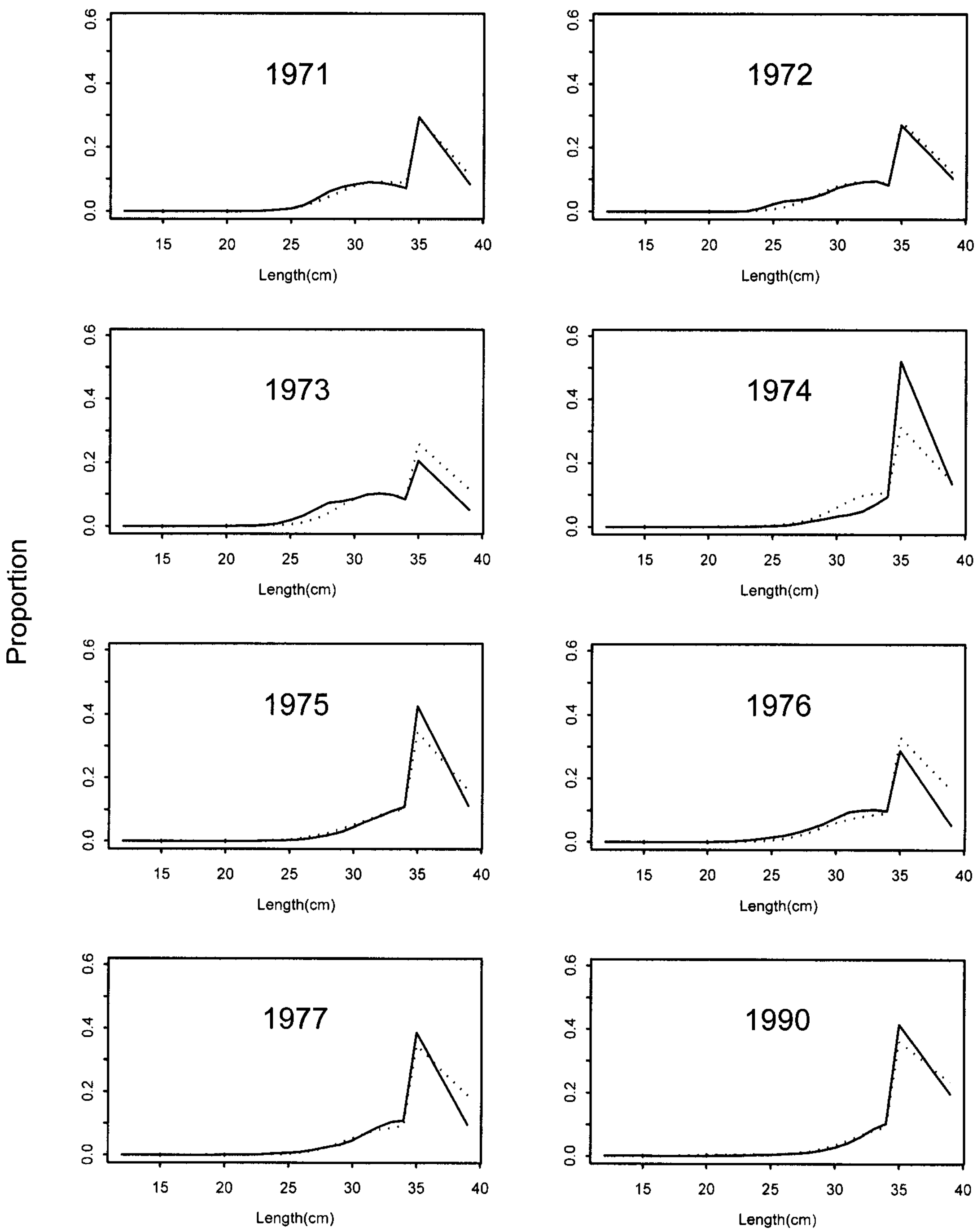


Figure 1.4 (continued). Fishery length compositions for GOA POP. Solid=observed, dashed=predicted.

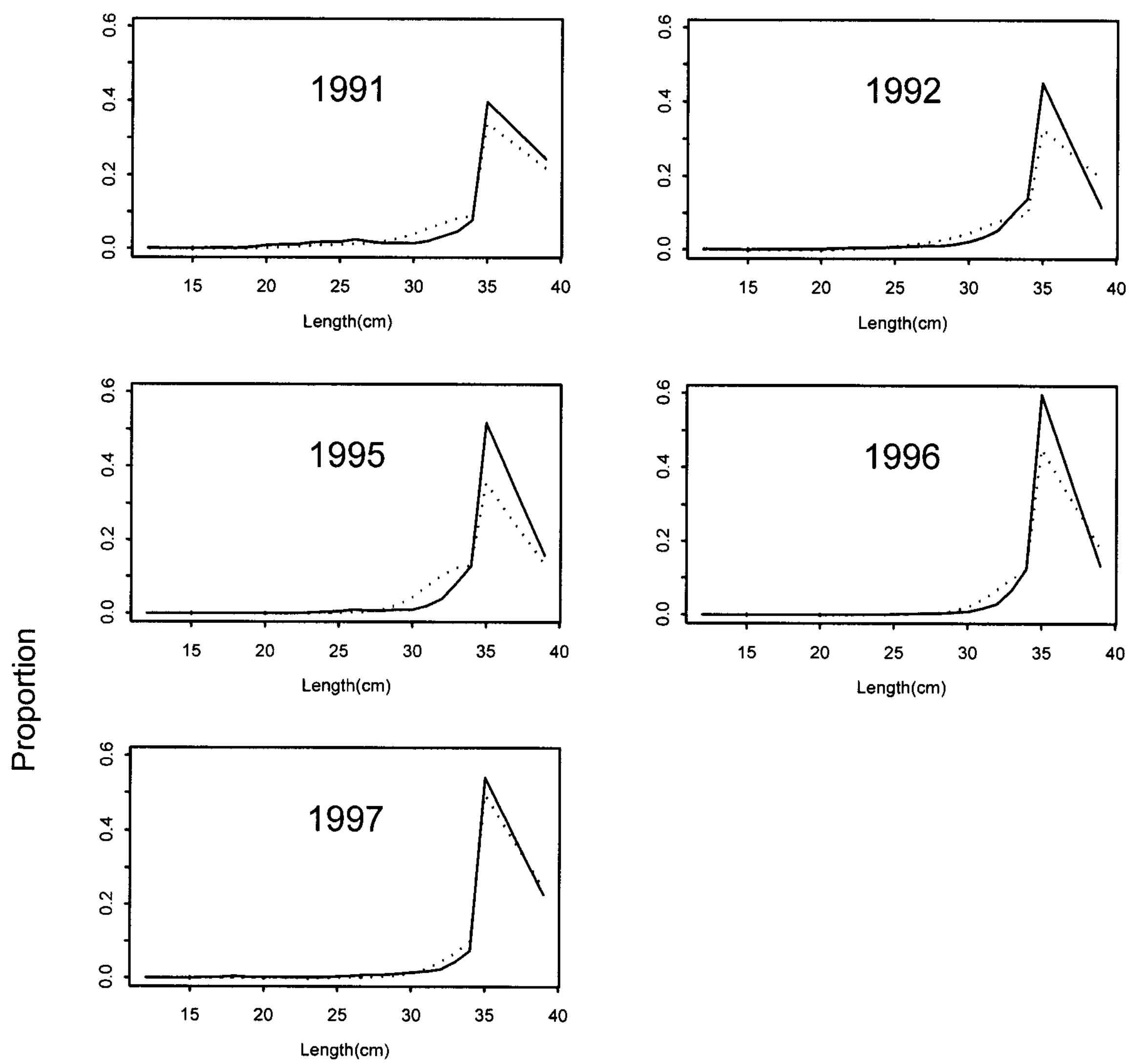


Figure 1.4 (continued). Fishery length composition for GOA POP. Observed=solid line, predicted=dotted line.

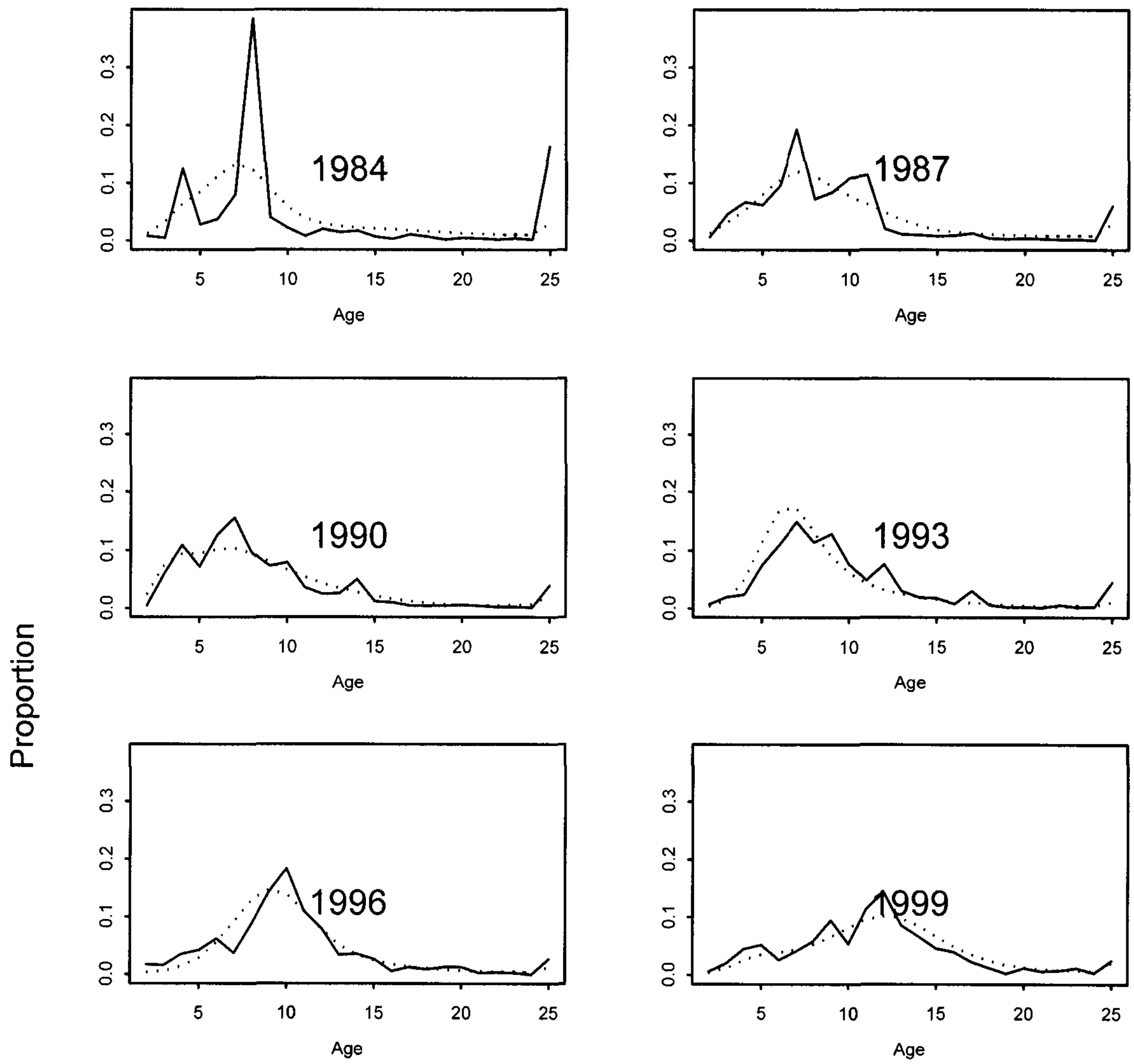


Figure 1.5. Survey age composition by year for GOA POP. Observed=solid line, predicted=dotted line.

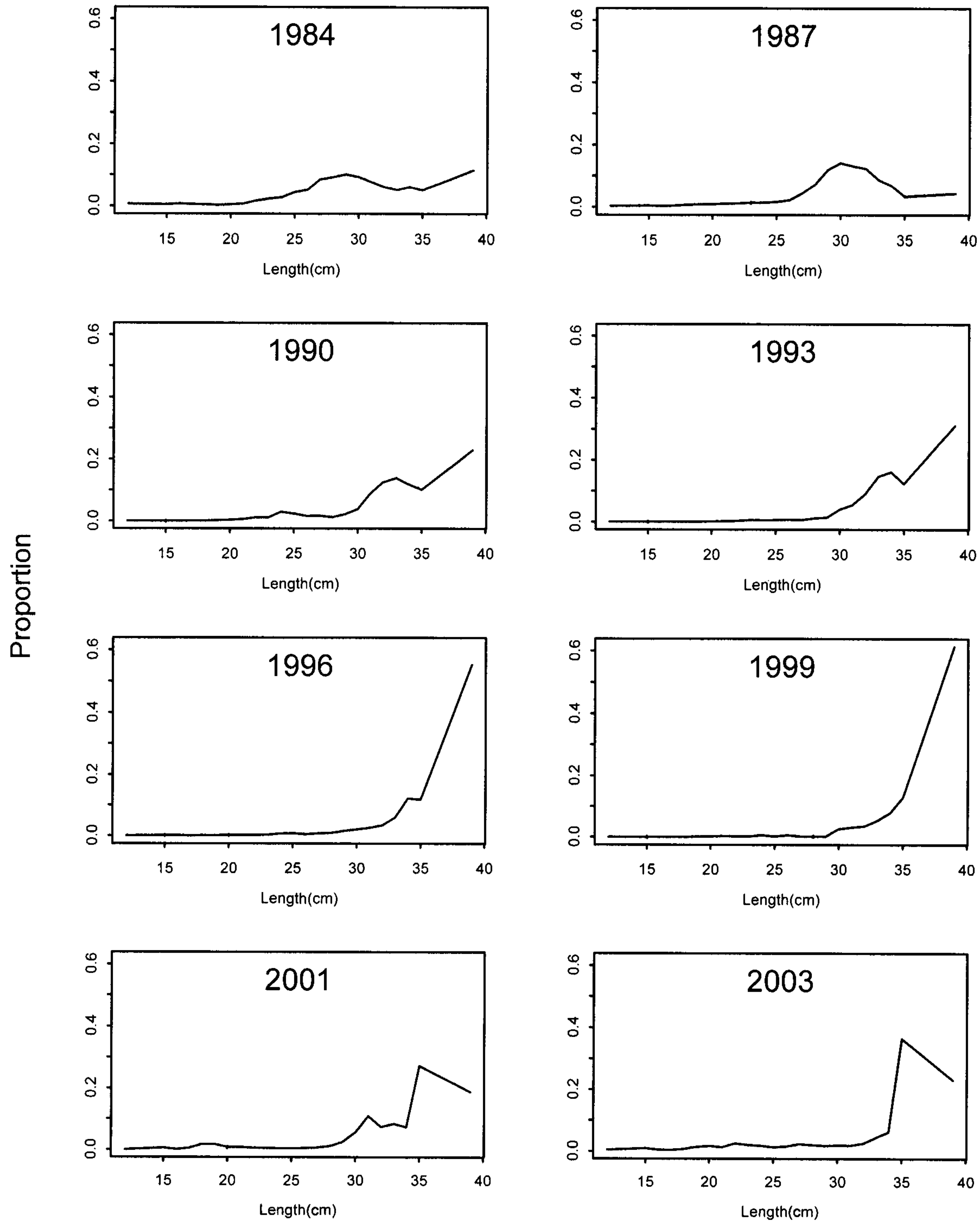


Figure 1.6. Survey length composition by year for GOA POP.

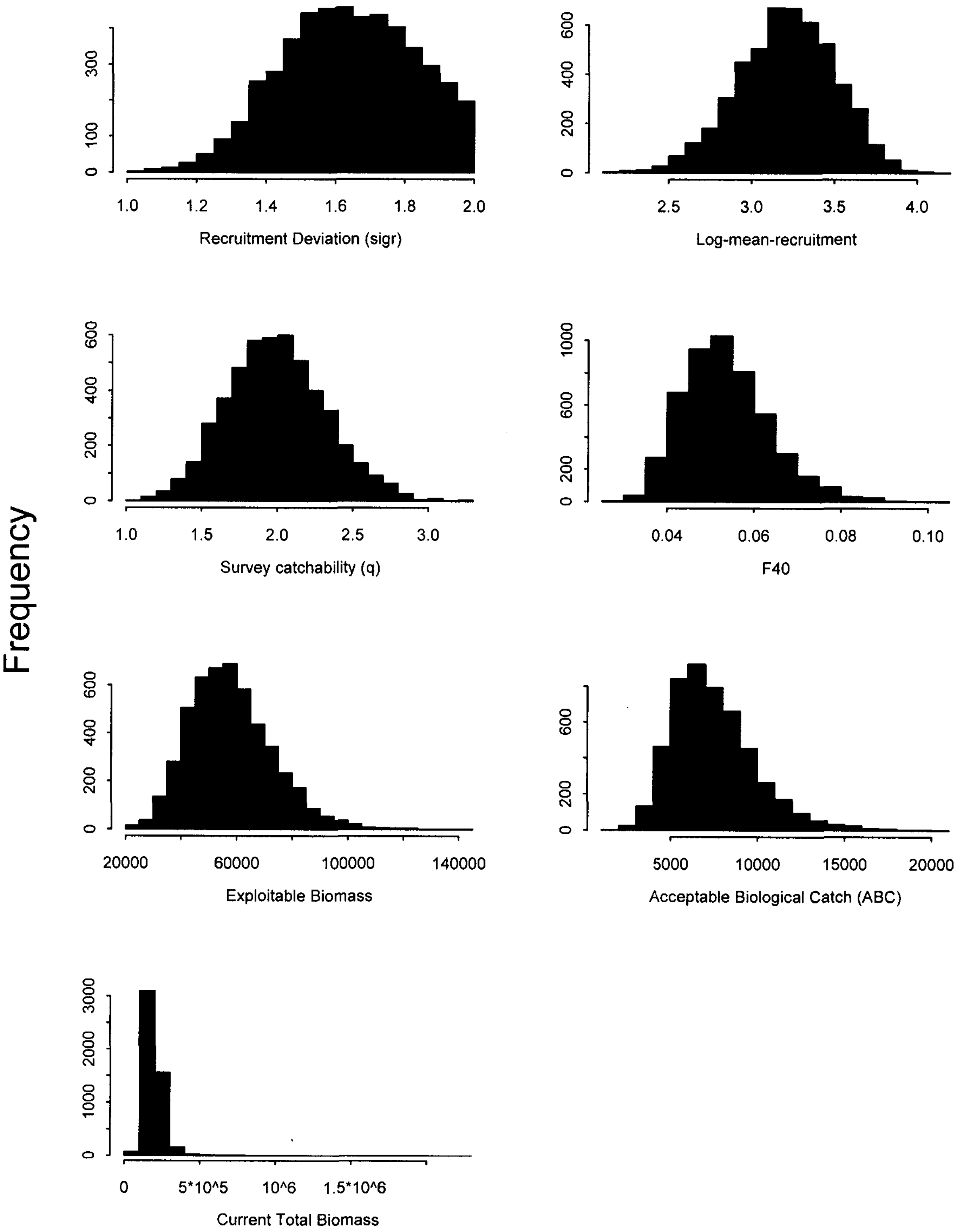


Figure 1.7. MCMC distributions of key parameters from a sample of 5 million runs for the base model.

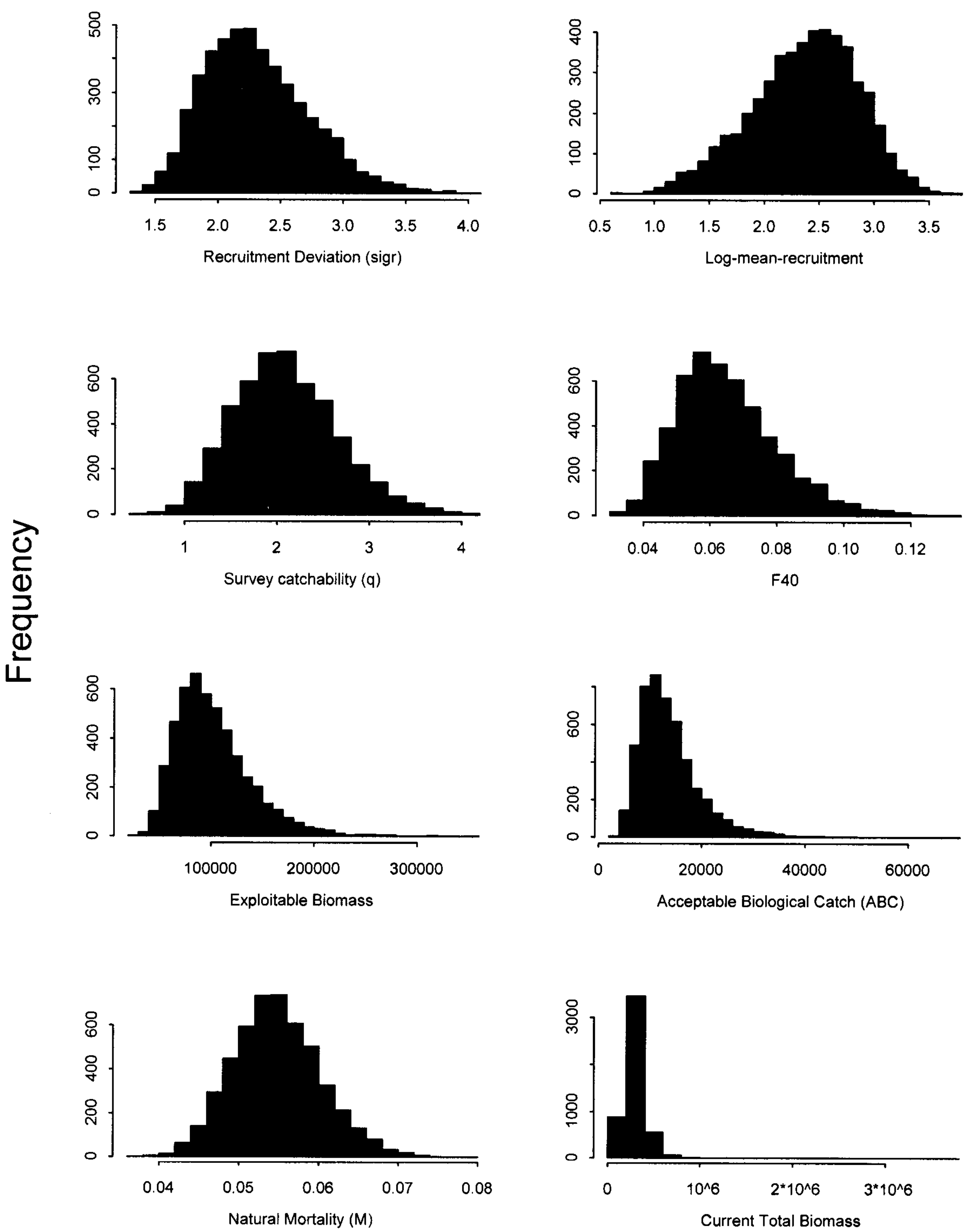


Figure 1.8. MCMC distributions of key parameters from a sample of 5 million runs for the recommended model.

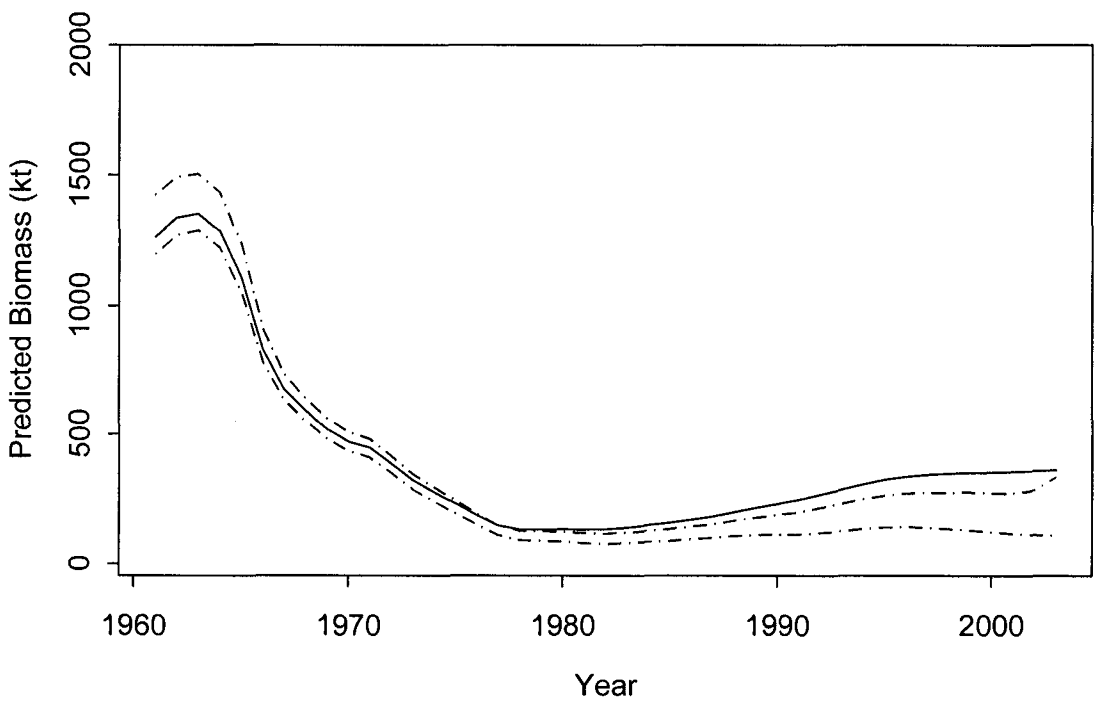


Figure 1.9. Time series of predicted total biomass for the base model. Dashed lines represent 95% confidence intervals from 5 million MCMC runs.

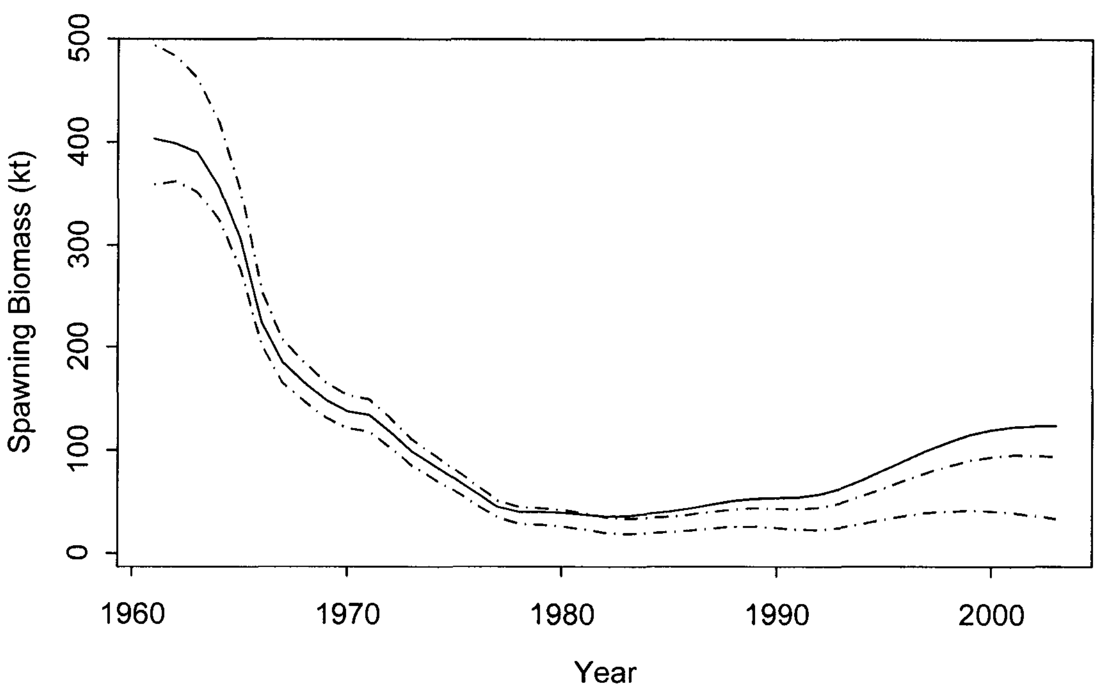


Figure 1.10. Time series of predicted spawning biomass of GOA POP for the base model. Dashed lines represent 95% confidence intervals from 5 million MCMC runs.

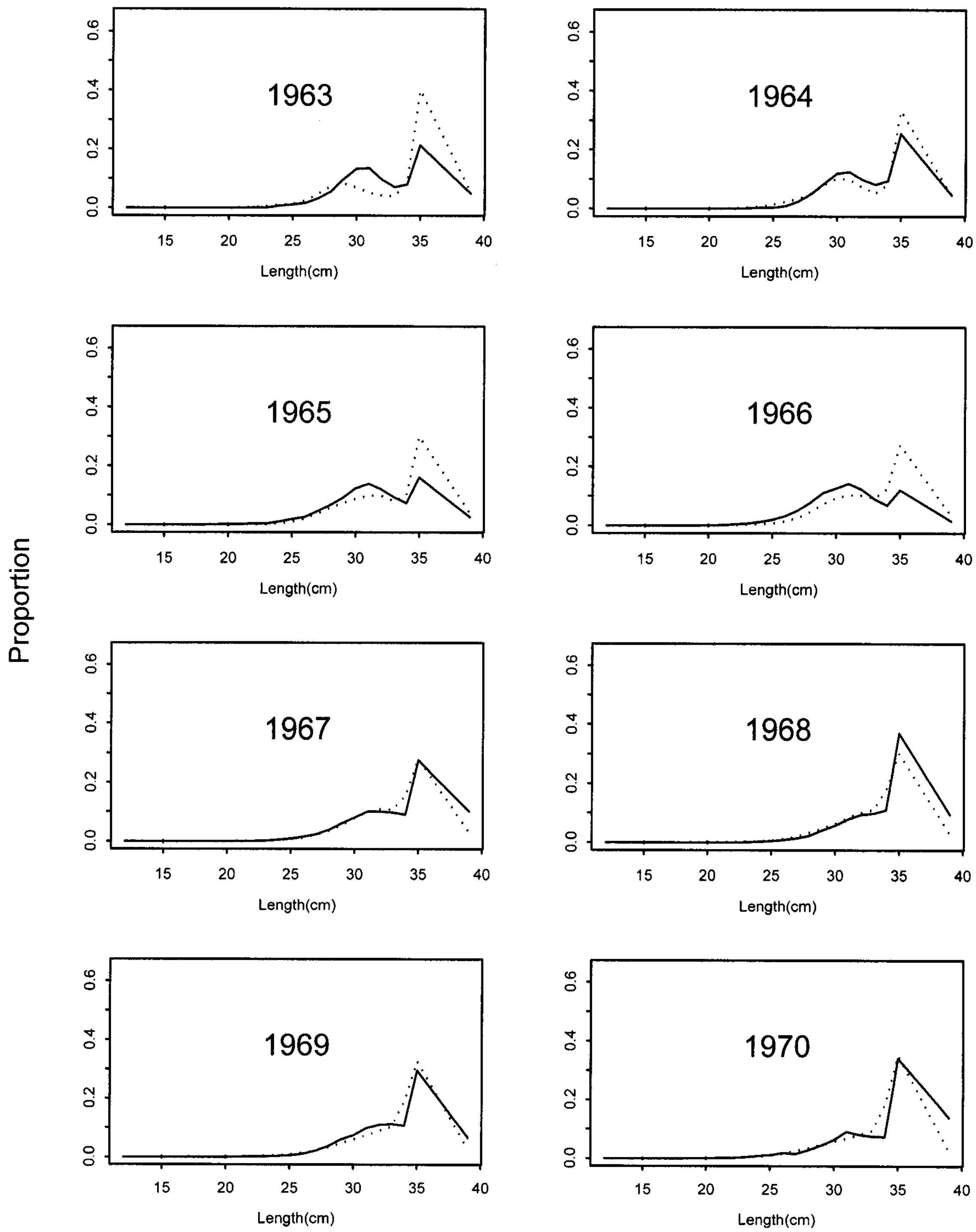


Figure 1.11. Fishery length compositions for GOA POP for base model. Observed=solid line, predicted=dotted line.

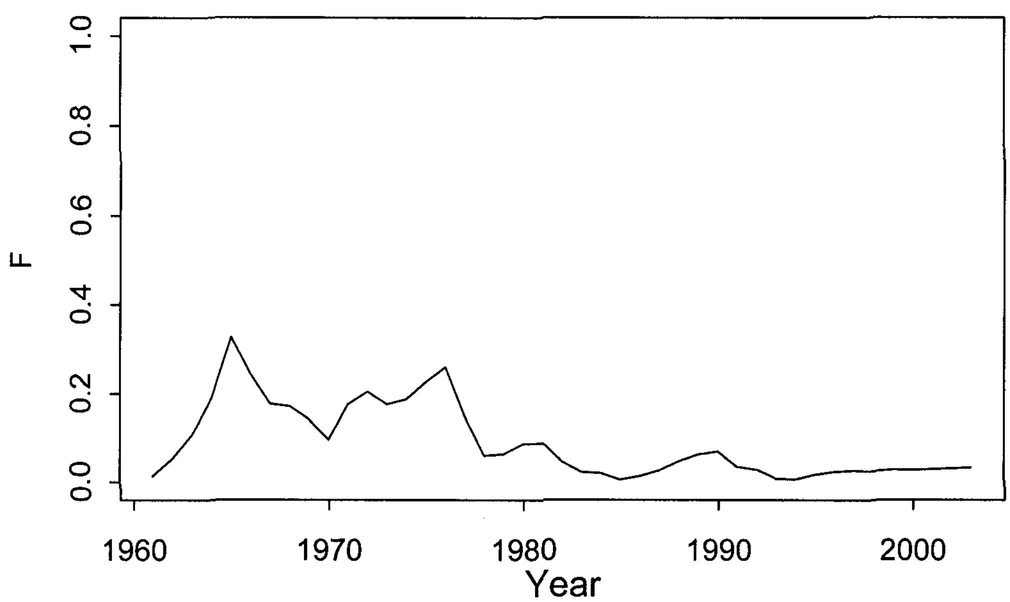


Figure 1.12. Time series of estimated fully selected fishing mortality for GOA POP from the base model.

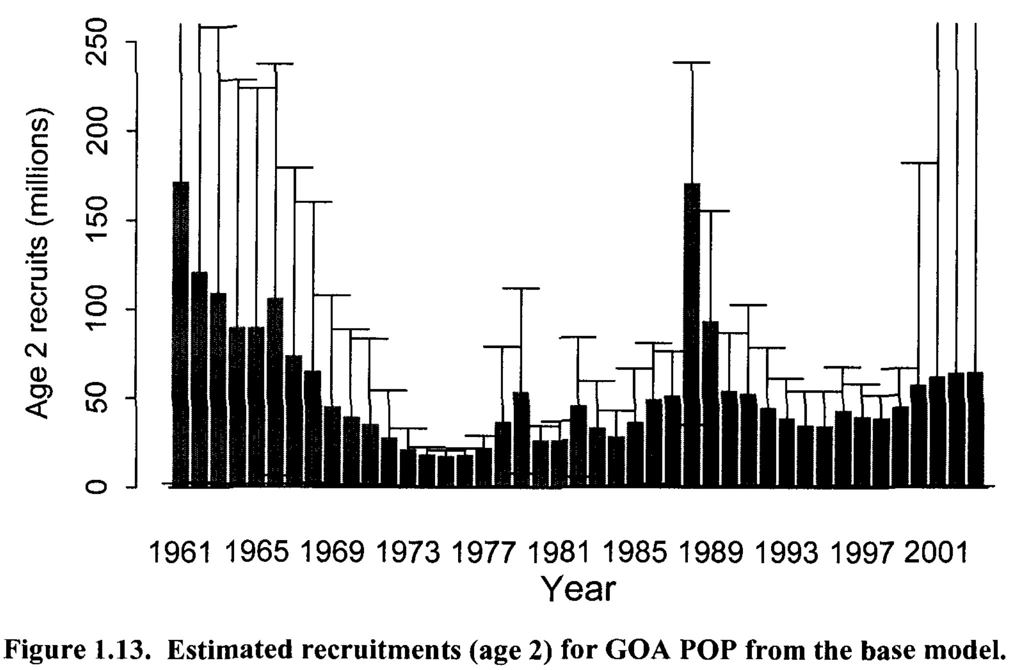


Figure 1.13. Estimated recruitments (age 2) for GOA POP from the base model.

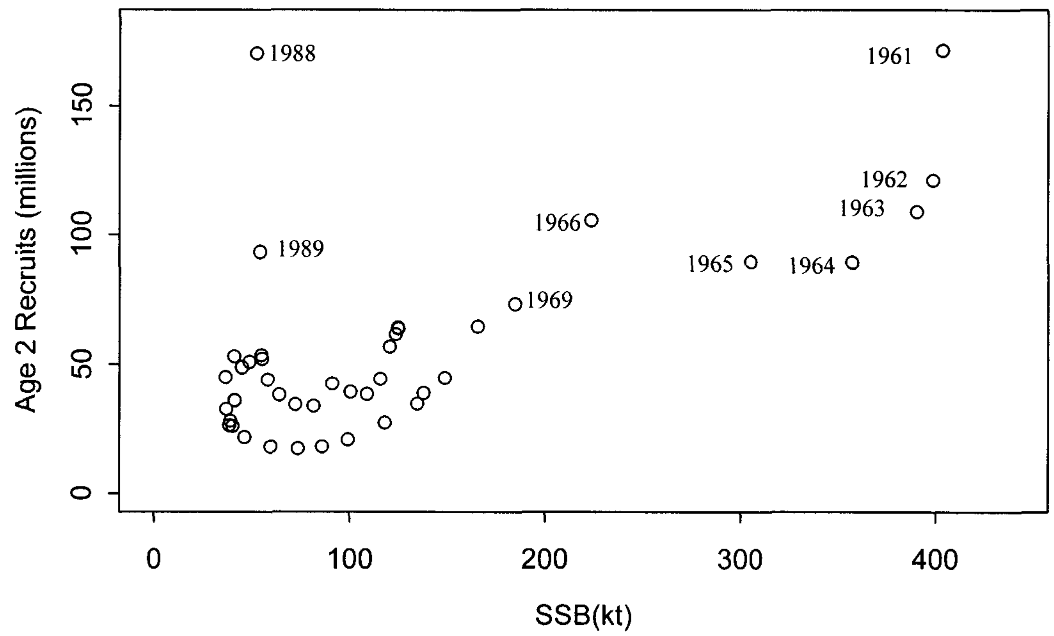


Figure 1.14. Scatterplot of spawner-recruit data for GOA POP estimated from the base model. Label is year class of age 2 recruits. SSB = Spawning stock biomass.

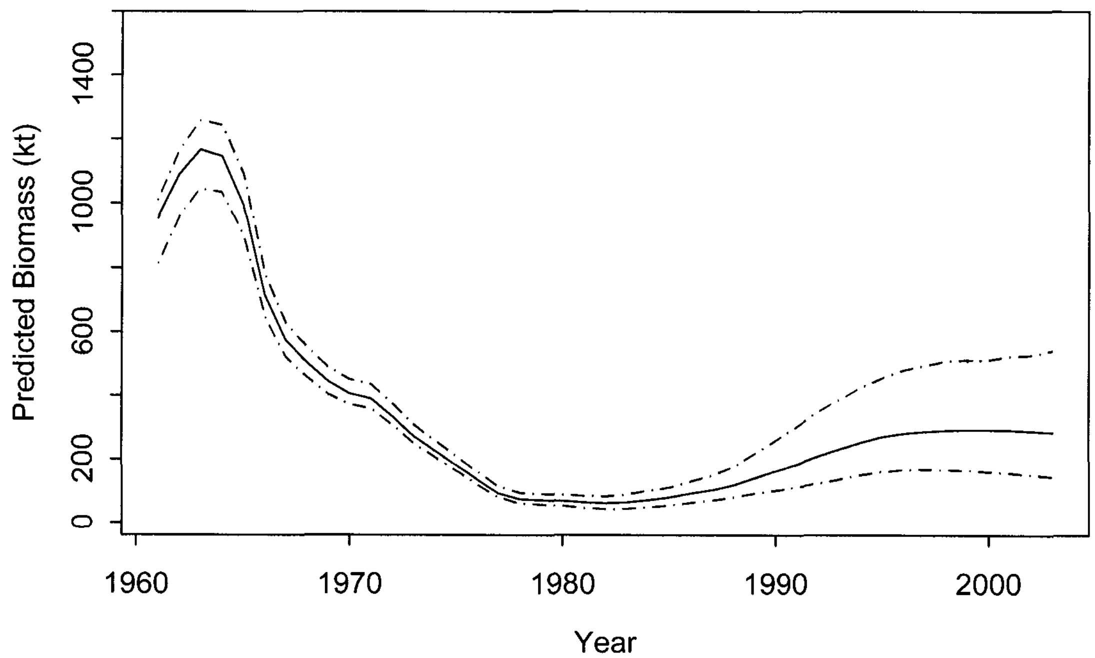


Figure 1.15. Time series of predicted total biomass of GOA POP from the recommended model. Dashed lines represent 95% confidence intervals from 5 million MCMC runs.

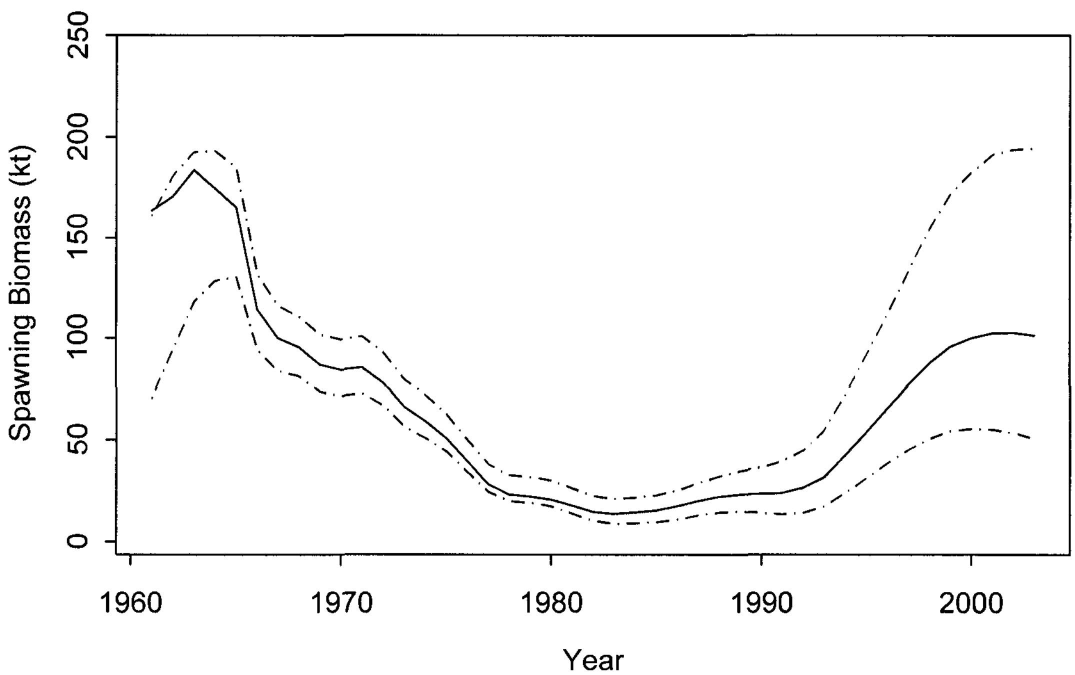


Figure 1.16. Time series of predicted spawning biomass of GOA POP from the recommended model. Dashed lines represent 95% confidence intervals from 5 million MCMC runs.

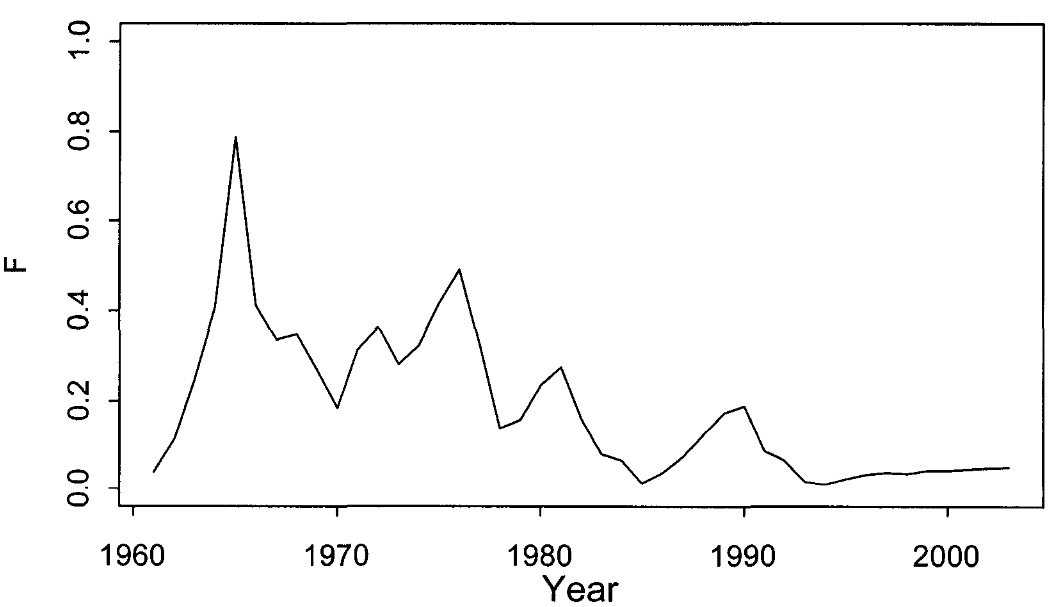


Figure 1.17. Time series of estimated fully selected fishing mortality for GOA POP from the recommended model.

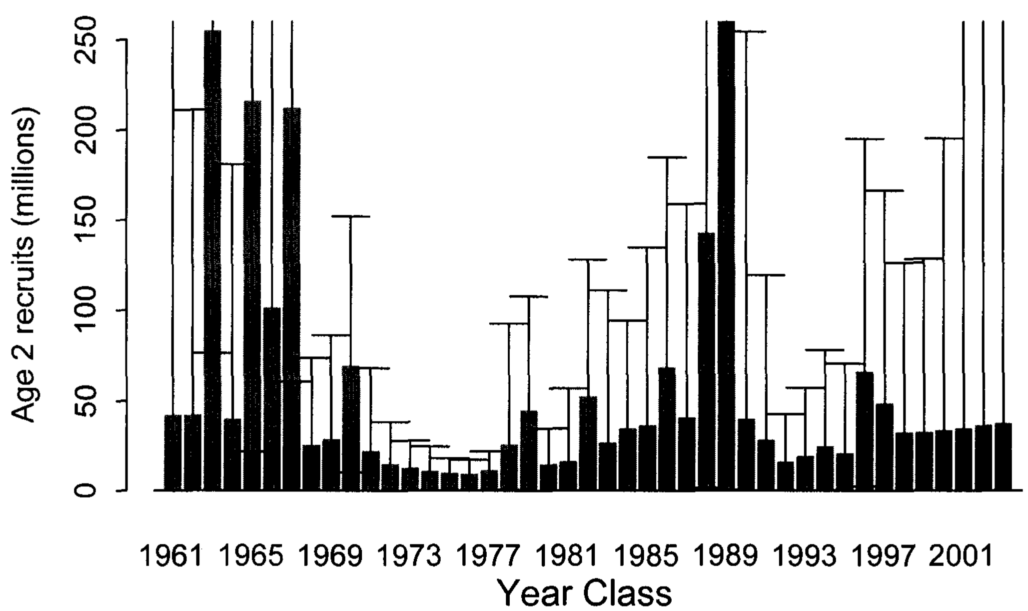


Figure 1.18. Times series of estimated recruitments (age 2) for GOA POP from the recommended model.

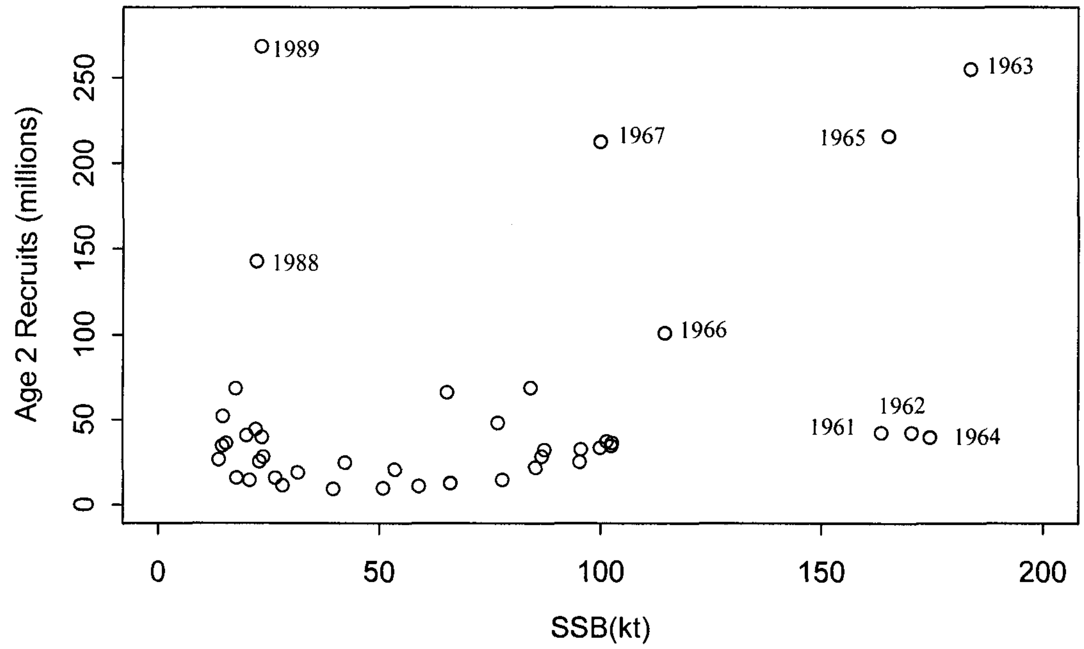


Figure 1.19. Scatterplot of spawner-recruit data for GOA POP estimated from the recommended model. Label is year class of age 2 recruits. SSB = Spawning stock biomass.

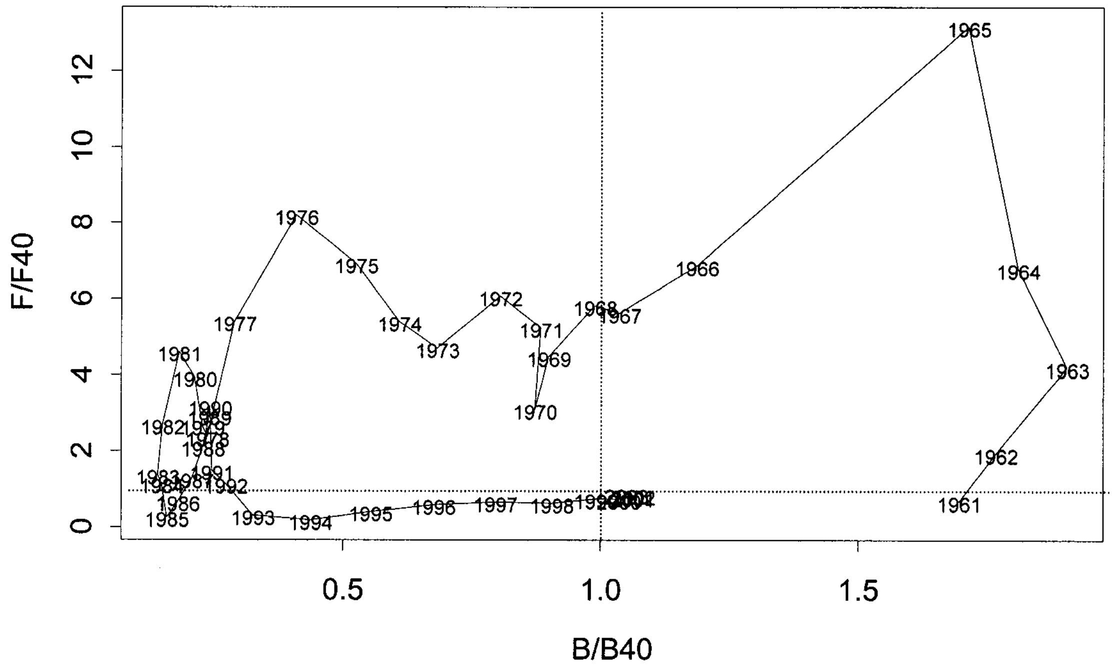


Figure 1.20. Time series of estimated fishing mortality over $F_{40\%}$ versus estimated spawning biomass over $B_{40\%}$.

Table 1.1a Commercial catch[a] (mt) of fish of Pacific ocean perch in the Gulf of Alaska, with Gulfwide values of acceptable biological catch (ABC) and fishing quotas[b] (mt), 1977-2002. Catches in 2003 updated through October 1, 2003.

		Regulatory Area				Gulfwide value	
Year	Fishery	Western	Central	Eastern	Total	ABC	Quota
1977	Foreign	6,282	6,166	10,993	23,441		
	U.S.	0	0	12	12		
	JV	-	-	-	-		
	Total	6,282	6,166	11,005	23,453	50,000	30,000
1978	Foreign	3,643	2,024	2,504	8,171		
	U.S.	0	0	5	5		
	JV	-	-	-	-		
	Total	3,643	2,024	2,509	8,176	50,000	25,000
1979	Foreign	944	2,371	6,434	9,749		
	U.S.	0	99	6	105		
	JV	1	31	35	67		
	Total	945	2,501	6,475	9,921	50,000	25,000
1980	Foreign	841	3,990	7,616	12,447		
	U.S.	0	2	2	4		
	JV	0	20	0	20		
	Total	841	4,012	7,618	12,471	50,000	25,000
1981	Foreign	1,233	4,268	6,675	12,176		
	U.S.	0	7	0	7		
	JV	1	0	0	1		
	Total	1,234	4,275	6,675	12,184	50,000	25,000
1982	Foreign	1,746	6,223	17	7,986		
	U.S.	0	2	0	2		
	JV	0	3	0	3		
	Total	1,746	6,228	17	7,991	50,000	11,475
1983	Foreign	671	4,726	18	5,415		
	U.S.	7	8	0	15		
	JV	1,934	41	0	1,975		
	Total	2,612	4,775	18	7,405	50,000	11,475
1984	Foreign	214	2,385	0	2,599		
	U.S.	116	0	3	119		
	JV	1,441	293	0	1,734		
	Total	1,771	2,678	3	4,452	50,000	11,475
1985	Foreign	6	2	0	8		
	U.S.	631	13	181	825		
	JV	211	43	0	254		
	Total	848	58	181	1,087	11,474	6,083
1986	Foreign	Tr	Tr	0	Tr		
	U.S.	642	394	1,908	2,944		

Table 1.1a (continued)

	JV	35	2	0	37		
	Total	677	396	1,908	2,981	10,500	3,702
1987	Foreign	0	0	0	0		
	U.S.	1,347	1,434	2,088	4,869		
	JV	108	4	0	112		
	Total	1,455	1,438	2,088	4,981	10,500	5,000
1988	Foreign	0	0	0	0		
	U.S.	2,586	6,467	4,718	13,771		
	JV	4	5	0	8		
	Total	2,590	6,471	4,718	13,779	16,800	16,800
1989	U.S.	4,339	8,315	6,348	19,002	20,000	20,000
1990	U.S.	5,203	9,973	5,938	21,114	17,700	17,700
1991	U.S.	1,589	2,956	2,087	6,631	5,800	5,800
1992	U.S.	1,266	2,658	2,234	6,159	5,730	5,200
1993	U.S.	477	1,140	443	2,060	3,378	2,560
1994	U.S.	165	920	768	1,853	3,030	2,550
1995	U.S.	1,422	2,598	1,722	5,742	6,530	5,630
1996	U.S.	987	5,145	2,246	8,378	8,060	6,959
1997	U.S.	1,832	6,720	979	9,531	12,990	9,190
1998	U.S.	850	7,501	610	8,961	12,820	10,776
1999	U.S.	1,935	7,910	627	10,472	13,120	12,590
2000	U.S.	1,160	8,379	618	10,157	13,020	13,020
2001	U.S.	944	9,249	624	10,817	13,510	13,510
2002	U.S.	2,720	8,261	748	11,729	13,190	13,190
2003	U.S.	2,073	7,848	606	10,627	13,663	13,660

Note: There were no foreign or joint venture catches after 1988. Catches prior to 1989 are landed catches only. Catches in 1989 and 1990 also include fish reported in weekly production reports as discarded by processors. Catches in 1991-2003 also include discarded fish, as determined through a "blend" of weekly production reports and information from the domestic observer program. Definitions of terms: JV = Joint venture; Tr = Trace catches; [a]Catch defined as follows: 1977, all *Sebastes* rockfish for Japanese catch, and Pacific ocean perch for catches of other nations; 1978, Pacific ocean perch only; 1979-87, the 5 species comprising the Pacific ocean perch complex; 1988-2003, Pacific ocean perch.[b]Quota defined as follows: 1977-86, optimum yield; 1987, target quota; 1988-2003 total allowable catch. Sources: Catch: 1977-84, Carlson et al. (1986); 1985-88, Pacific Fishery Information Network (PacFIN), Pacific Marine Fisheries Commission, 305 State Office Building, 1400 S.W. 5th Avenue, Portland, OR 97201; 1989-2003, National Marine Fisheries Service, Alaska Region, P.O. Box 21668, Juneau, AK 99802. ABC and Quota: 1977-1986 Karinen and Wing (1987); 1987-2000, Heifetz et al. (2000); 2001-2003, Heifetz et al. (2002).

Table 1.1b. Catch (mt) of Pacific ocean perch taken in research cruises in the Gulf of Alaska, 1977-2003. (Does not include longline survey catch before 1995; tr=trace).

Year	Catch
1977	13.0
1978	5.7
1979	12.2
1980	12.6
1981	57.1
1982	15.2
1983	2.4
1984	76.5
1985	35.2
1986	14.4
1987	68.8
1988	0.3
1989	1.0
1990	25.5
1991	0.1
1992	0.0
1993	59.2
1994	tr
1995	tr
1996	81.2
1997	tr
1998	305.0
1999	330.2
2000	0.0
2001	42.5
2002	tr
2003	50.4

Table 1.2 Fishery length frequency data for Pacific ocean perch in the Gulf of Alaska.

Length Class (cm)	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003
<13	5	0	14	0	0	1	0	34	0	0	0	0	0	0
13-15	26	11	22	3	0	3	2	11	1	1	2	1	1	0
16	13	16	16	2	0	1	0	23	0	1	2	0	1	0
17	19	13	13	2	0	2	0	35	2	1	3	0	1	1
18	31	17	13	6	0	2	6	69	2	3	2	7	1	1
19	46	26	20	9	2	3	5	25	3	4	1	7	3	1
20	72	38	23	20	3	4	6	25	12	3	3	8	7	0
21	124	37	32	35	2	5	7	27	19	5	14	9	16	2
22	177	50	54	60	9	7	11	30	21	11	14	15	21	3
23	235	66	81	96	19	18	22	37	17	13	15	15	29	5
24	321	81	112	129	31	20	25	34	44	30	30	15	33	6
25	412	97	167	166	64	34	44	53	61	37	24	26	50	10
26	512	123	239	198	85	56	83	89	90	47	45	23	71	12
27	642	158	303	250	97	80	158	143	88	44	70	41	83	22
28	724	156	338	315	125	110	272	191	117	40	80	49	123	30
29	836	240	416	359	137	158	427	287	201	94	92	66	135	36
30	951	263	496	398	167	174	666	499	312	83	101	92	133	47
31	1089	319	531	440	179	225	948	855	516	147	160	114	207	62
32	1259	382	584	472	192	254	1443	1312	860	271	229	176	234	71
33	1374	439	644	490	212	283	2353	1995	1420	463	340	320	404	102
34	1418	485	674	523	216	306	3646	2508	2338	739	665	479	671	194
35-38	5601	1918	2477	1767	746	1158	17318	14246	17214	4967	5785	4390	4689	1755
>38	4249	1567	2019	1051	563	783	5329	5554	7481	2250	3016	2612	2992	1242
Total	20136	6502	9288	6791	2849	3687	32771	28082	30819	9254	10693	8465	9905	3602

Table 1.3. Fishery age compositions for GOA Pacific ocean perch 1998-2002.

			Year		
Age Class	1998	1999	2000	2001	2002
2	0.001	-	-	-	-
3	-	-	-	0.004	-
4	0.002	-	0.008	0.003	0.002
5	0.001	0.002	0.014	0.004	0.008
6	0.001	0.014	0.029	0.011	0.011
7	0.005	0.024	0.018	0.029	0.029
8	0.031	0.045	0.046	0.025	0.085
9	0.076	0.045	0.051	0.051	0.072
10	0.180	0.054	0.063	0.041	0.106
11	0.122	0.173	0.066	0.052	0.091
12	0.132	0.189	0.130	0.075	0.058
13	0.106	0.128	0.103	0.139	0.071
14	0.120	0.090	0.095	0.112	0.114
15	0.052	0.116	0.102	0.088	0.111
16	0.029	0.054	0.079	0.086	0.071
17	0.051	0.019	0.050	0.069	0.058
18	0.020	0.021	0.040	0.071	0.042
19	0.014	0.002	0.030	0.046	0.032
20	0.008	0.002	0.012	0.019	0.014
21	0.011	-	0.017	0.019	0.008
22	0.004	0.009	0.014	0.006	0.006
23	0.008	0.002	0.006	0.012	0.003
24	0.003	-	0.003	0.006	0.002
25	0.022	0.009	0.024	0.032	0.008
Sample size	1336	423	1312	1234	624

Table 1.4. Biomass estimates (mt) and Gulfwide confidence intervals for Pacific ocean perch in the Gulf of Alaska based on the 1984-2003 trawl surveys. (Biomass estimates and confidence intervals for 2001 have been slightly revised from those listed in previous SAFE reports for slope rockfish.)

	Western	Central		Eastern			
	Shumagin	Chirikof	Kodiak	Yakutat	Southeast	Total	95% Confidence interval
1984	59,710	9,672	36,976	94,055	32,280	232,694	101,550 - 363,838
1987	62,906	19,666	44,441	35,612	52,201	214,827	125,499 - 304,155
1990	24,375	15,991	15,221	35,635	46,780	138,003	70,993 - 205,013
1993	75,416	103,224	153,262	50,048	101,532	483,482	260,553 - 706,411
1996	92,618	140,479	326,280	50,394	161,641	771,413	355,756 - 1,187,069
1999	38,196	402,293	209,675	32,733	44,367	727,263	0 - 1,566,566
2001*	275,210	39,819	385,126	44,392	102,514	820,061	364,570 – 1,275,552
2003	72,851	116,231	166,815	27,762	73,737	457,394	313,363 – 601,426

*The 2001 survey did not sample the eastern Gulf of Alaska (the Yakutat and Southeastern areas). Substitute estimates of biomass for the Yakutat and Southeastern areas were obtained by averaging the biomass estimates for Pacific ocean perch in these areas in the 1993, 1996, and 1999 surveys, that portion of the variance was obtained by using a weighted average of the three prior surveys' variance. Confidence intervals that contain zero are a result of using a normal approximation.

Table 1.5. Survey age composition (% frequency) data for Pacific ocean perch in the Gulf of Alaska. Age compositions for are based on "break and burn" reading of otoliths.

Age	1984	1987	1990	1993	1996	1999
2	0.007	0.009	0.014	0.027	0.010	0.046
3	0.002	0.085	0.059	0.046	0.031	0.099
4	0.061	0.101	0.116	0.050	0.063	0.099
5	0.029	0.058	0.095	0.071	0.070	0.111
6	0.052	0.061	0.114	0.102	0.111	0.060
7	0.115	0.115	0.097	0.102	0.058	0.061
8	0.386	0.047	0.073	0.090	0.075	0.058
9	0.028	0.056	0.063	0.114	0.111	0.065
10	0.016	0.084	0.058	0.064	0.130	0.030
11	0.007	0.104	0.037	0.034	0.077	0.058
12	0.013	0.021	0.025	0.039	0.058	0.072
13	0.010	0.013	0.026	0.032	0.025	0.040
14	0.012	0.012	0.070	0.020	0.022	0.036
15	0.005	0.012	0.015	0.029	0.019	0.021
16	0.003	0.016	0.012	0.013	0.007	0.025
17	0.008	0.018	0.006	0.044	0.015	0.012
18	0.005	0.010	0.008	0.010	0.011	0.009
19	0.002	0.006	0.006	0.003	0.018	0.003
20	-	0.009	0.007	0.003	0.017	0.008
21	0.004	0.007	0.007	0.003	0.007	0.005
22	0.003	0.003	0.002	0.005	0.006	0.009
23	0.002	0.004	0.003	0.003	0.003	0.014
24	0.006	0.003	0.005	0.005	-	0.005
25	0.224	0.147	0.083	0.091	0.056	0.052
Total	2575	1824	1788	1492	718	963

Table 1.6. Estimated numbers (thousands) in 2003, fishery selectivity, and survey selectivity of Pacific ocean perch in the Gulf of Alaska. Also shown are schedules of age specific weight and female maturity.

Age	Numbers in 2003 (1000's)	Percent mature	Weight (g)	Fishery selectivity	Survey selectivity
2	37,024	0	46	0	2
3	34,159	0	106	1	6
4	30,592	0	180	2	18
5	27,904	0	261	3	33
6	25,496	0	342	8	48
7	23,834	12	420	29	97
8	33,051	20	493	100	100
9	40,633	30	559	95	100
10	11,453	42	619	95	100
11	12,336	56	672	95	100
12	8,528	69	718	95	100
13	6,483	79	758	95	100
14	10,497	87	792	95	100
15	13,564	92	822	95	100
16	84,867	95	847	95	100
17	41,663	97	868	95	100
18	10,839	98	886	95	100
19	15,793	99	902	95	100
20	7,027	99	915	95	100
21	5,308	100	926	95	100
22	3,324	100	935	95	100
23	5,557	100	943	95	100
24	1,492	100	950	95	100
25+	10,702	100	970	95	100

Table 1.7. Summary of results from five alternative Pacific ocean perch models

	Base Model		Model 2		Model 3		Model 4		Model 5	
Likelihoods	**Value**	**Weight**	**Value**	**Weight**	**Value**	**Weight**	**Value**	**Weight**	**Value**	**Weight**
Catch	1.71	50	0.17	50	0.10	50	0.16	50	0.09	50
Survey Biomass	9.34	1	7.42	1	6.72	1	11.24	1	6.82	1
Fishery Ages	53.79	1	37.32	1	32.10	1	35.13	1	32.87	1
Survey Ages	77.58	1	80.77	1	67.76	1	69.55	1	67.59	1
Fishery Sizes	213.61	1	62.71	1	54.40	1	58.68	1	50.82	1
Data-Likelihood	356.03		188.39		161.08		174.76		158.20	
Penalties/Priors										
Recruitment Devs	7.93	50	31.03	1	31.53	1	35.69	1	32.50	1
Fishery Selectivity	4.60	12.5	2.40	1	2.34	1	1.37	1	1.92	1
Survey Selectivity	1.72	12.5	1.48	1	0.92	1	0.71	1	0.84	1
Fish-Sel Dome	0.06	1,000	0.00	1	0.04	1	0.00	1	0.00	1
Survey-Sel Dome	0.19	1,000	0.00	1	0.01	1	0.00	1	0.00	1
Average Selectivity	0.00	10	0.00	1	0.00	1	0.00	1	0.00	1
F Regularity	49.26	1	7.28	0.1	4.40	0.1	12.26	0.2	4.74	0.1
σ_r prior	0.17		0.71		0.69		0.01		0.65	
q prior	0.10		0.27		1.82		0.02		0.99	
Objective Fun Total										
-ln$\mathcal{L}$ (- log likelihood)	420.07		231.56		202.83		224.81		199.84	
		LN Prior		LN Prior		LN Prior		LN Prior		LN Prior
Parameter Ests.		(μ,σ)		(μ,σ)		(μ,σ)		(μ,σ)		(μ,σ)
q	1.22	(1,.2)	1.39	(1,0.2)	2.35	(1,0.2)	1.00	(1, 1e-5)	1.88	(1,0.2)
M	0.05	Fixed	0.06	(.05,.01)	0.05	Fixed	0.05	Fixed	0.06	(.05,.01)
σ_r	0.69	(.9,.2)	1.00	(1.7,.2)	1.01	(1.7,.2)	1.05	(1.7,.2)	1.02	(1.7,.2)
log-mean-rec	4.17		3.82		3.38		3.61		3.61	
F_{40}	0.05		0.06		0.05		0.05		0.06	
Total Biomass	360,650		384,060		250,510		508,230		285,070	
B_{2004}	120,090		138,385		85,840		166,100		95,762	
B_0	280,254		290,955		238,918		366,406		224,248	
B_{40}	***112,102***		***116,382***		***95,567***		***146,562***		***89,699***	
ABC_{F40}	***14,761***		***18,519***		***9,406***		***19,877***		***13,336***	
F_{50}	0.04		0.04		0.04		0.04		0.04	
ABC_{F50}	10,405		13,132		6,608		13,958		9,410	

Table 1.8. Estimates of key parameters with MLE estimates of standard error and 95% Bayesian confidence intervals (BCI) derived from MCMC simulations.

Parameter	μ	σ	σ(MCMC)	BCI-Lower	BCI-Upper
q	1.88	0.508	0.560	1.127	3.330
M	0.059	0.006	0.005	0.045	0.066
$F_{40\%}$	0.060	0.015	0.015	0.042	0.100
B_{2003}	101,380	32,465	37,843	50,462	193,829
ABC	13,363	4,732	5,923	5,713	28298
σ_r	1.02	0.114	0.419	1.64	3.25

Table 1.9. Estimated time series of female spawning biomass, 6+ biomass (age 6 and greater), catch/6 + biomass, and number of age two recruits for Pacific ocean perch in the Gulf of Alaska. Estimates are shown for the current assessment and from the previous SAFE.

	Spawning biomass (mt)		6+ Biomass (mt)		Catch/6+ biomass		Age 2 recruits (1000's)	
Year	Current	Previous	Current	Previous	Current	Previous	Current	Previous
1977	39,481	48,907	130,740	141,950	0.348	0.152	9,169	22,517
1978	28,010	43,760	88,930	125,765	0.243	0.064	11,166	38,014
1979	22,888	43,432	70,370	123,057	0.114	0.068	25,273	60,901
1980	22,029	42,761	65,106	120,099	0.127	0.091	44,380	26,518
1981	20,540	40,900	59,212	115,668	0.182	0.092	14,192	26,667
1982	17,619	39,041	51,184	116,290	0.205	0.047	15,698	47,485
1983	14,558	39,482	47,324	129,758	0.114	0.022	51,844	34,679
1984	13,679	41,747	54,578	137,442	0.052	0.021	26,420	29,138
1985	14,429	44,259	57,716	144,924	0.048	0.006	34,333	36,849
1986	15,374	47,989	61,049	160,221	0.013	0.014	35,957	49,830
1987	17,390	52,157	75,919	170,829	0.029	0.027	68,177	50,326
1988	19,972	55,888	83,767	177,172	0.054	0.049	40,740	159,199
1989	22,075	58,139	91,216	181,092	0.094	0.066	142,264	80,177
1990	23,249	59,069	94,916	185,422	0.124	0.07	268,167	45,291
1991	23,379	59,642	103,941	189,242	0.126	0.035	39,487	42,186
1992	23,797	62,590	105,798	232,133	0.063	0.027	27,923	36,468
1993	26,344	68,233	141,489	257,436	0.044	0.008	15,680	32,123
1994	31,459	76,017	215,674	277,060	0.01	0.007	18,759	28,812
1995	42,237	84,788	241,154	294,249	0.008	0.02	24,560	26,614
1996	53,458	93,210	260,376	303,904	0.022	0.028	20,615	33,679
1997	65,267	101,074	268,391	307,698	0.031	0.031	65,941	42,751
1998	76,809	107,773	270,578	307,613	0.035	0.03	48,124	43,633
1999	87,391	112,964	270,082	305,958	0.033	0.035	32,253	47,125
2000	95,610	115,830	266,981	303,597	0.039	0.034	32,396	60,147
2001	99,941	117,186	272,777	303,634	0.037	0.036	33,361	62,901
2002	102,503	117,090	274,761	303,281	0.039	0.039	34,449	63,966
2003	102,644	112,269	271,652	298,816	0.043		36,246	47,840
2004*	95,760		266,963				37,024	

Table 1.10. Allocation of ABC for 2004 Pacific ocean perch in the Gulf of Alaska.

		Western	Central		Eastern		
Year	Weights	Shumagin	Chirikof	Kodiak	Yakutat	Southeast	Total
1999	4	5%	55%	29%	5%	6%	100%
2001	6	32%	5%	45%	5%	12%	100%
2003	9	16%	25%	36%	6%	16%	100%
Weighted Mean	19	19%	25%	38%	5%	13%	100%
Allocation		19%	63%		18%		
Area ABC		2,522	8,384		2,430		13,336

Table 1.11. Bycatch (kg) and bycatch rates during 1997 - 2002 of living substrates in the Gulf of Alaska for combined rockfish fisheries, all gears. Source: Gaichas and Ianelli, unpublished data.

	1997	1998	1999	2000	2001	2002	Average
Non-target species	Bycatch (kg)						
Sea Pens/Whips	0	0	23	12	30	18	14
Sponges	1,504	643	5,393	1,482	1,887	1,951	2,143
Anemones	459	15	673	1,438	255	335	529
Tunicates	14	45	6	481	8	38	99
Echinoderms	2,023	532	2,016	773	2,952	683	1,496
Coral	1,636	330	766	10,005	4,317	15,143	5,366
Rockfish Catch (tons)	13,083	13,592	18,333	15,947	15,672	16,977	15,601
	Bycatch rate (kg/mt target)						
Sea Pens/Whips	0.0000	0.0000	0.0012	0.0007	0.0019	0.0010	0.0009
Sponges	0.1150	0.0473	0.2941	0.0929	0.1204	0.1149	0.1374
Anemones	0.0351	0.0011	0.0367	0.0902	0.0163	0.0197	0.0339
Tunicates	0.0011	0.0033	0.0003	0.0301	0.0005	0.0022	0.0063
Echinoderms	0.1546	0.0391	0.1099	0.0485	0.1883	0.0402	0.0959
Coral	0.1251	0.0242	0.0418	0.6274	0.2755	0.8920	0.3440

1.9 References

Ackley, D. R. and J. Heifetz. 2001. Fishing practices under maximum retainable bycatch rates in Alaska's groundfish fisheries. Alaska Fish. Res. Bull. 8:22-44.

Archibald, C. P., W. Shaw, and B. M. Leaman. 1981. Growth and mortality estimates of rockfishes (Scorpaenidae) from B.C. coastal waters, 1977-1979. Can. Tech. Rep. Fish. Aquat. Sci. 1048: iv +57 p.

Byerly, Michael M. 2001. The ecology of age-1 Copper Rockfish (*Sebastes caurinus*) in vegetated habitats of Sitka sound, Alaska. M.S. thesis. University of Alaska, Fairbanks. Fisheries Division, 11120 Glacier Hwy, Juneau, AK 99801.

Carlson, H.R., D.H. Ito, R.E. Haight, T.L. Rutecki, and J.F. Karinen. 1986. Pacific ocean perch. In R.L. Major (editor), Condition of groundfish resources of the Gulf of Alaska region as assessed in 1985, p. 155-209. U.S. Dept. Commerce, NOAA Tech. Memo. NMFS F/NWC-106.

Chilton, D.E. and R.J. Beamish. 1982. Age determination methods for fishes studied by the groundfish program at the Pacific Biological Station. Can. Spec. Pub. Fish. Aquat. Sci. 60.

Courtney, D.L., J. Heifetz, M. F. Sigler, and D. M. Clausen. 1999. An age structured model of northern rockfish, *Sebastes polyspinis*, recruitment and biomass in the Gulf of Alaska. In Stock assessment and fishery evaluation report for the groundfish resources of the Gulf of Alaska as projected for 2000. North Pacific Fishery Management Council, 605 W 4th Ave, Suite 306 Anchorage, AK 99501. pp. 361-404.

Dorn, M. W. 2002. Advice on west coast rockfish harvest rates from Bayesian meta-analysis of stock-recruit relationships. North Amer. J. Fish. Mgmt. 22:280-300.

Gelman, A., J.B. Carlin, H.S. Stern and D.B. Rubin. 1995. Bayesian data analysis. Chapman and Hall, London. 526 pp.

Gharrett, A. J., A.K. Gray, and J. Heifetz. 2001. Identification of rockfish (*Sebastes* spp.) from restriction site analysis of the mitochondrial NM-3/ND-4 and 12S/16S rRNA gene regions. Fish. Bull. 99:49-62.

Goodman, D., M. Mangel, G. Parkes, T.J. Quinn II, V. Restrepo, T. Smith, and K. Stokes. 2002. Scientific Review of the Harvest Strategy Currently Used in the BSAI and GOA Groundfish Fishery Management Plans. Draft report. North Pacific Fishery Management Council, 605 W 4th Ave, Suite 306 Anchorage, AK 99501. 153 pp.

Gunderson, D. R. 1977. Population biology of Pacific Ocean perch, *Sebastes alutus*, stocks in the Washington-Queen Charlotte Sound region, and their response to fishing. Fish. Bull. 75(2):369–403.

Haldorson, L, and M. Love. 1991. Maturity and fecundity in the rockfishes, *Sebastes spp.*, a review. Mar. Fish. Rev. 53(2):25–31.

Hanselman, D.H., T.J. Quinn II, C. Lunsford, J. Heifetz and D.M. Clausen. 2001. Spatial implications of adaptive cluster sampling on Gulf of Alaska rockfish. In Proceedings of the 17th Lowell-Wakefield Symposium: Spatial Processes and Management of Marine Populations, pp. 303-325. Univ. Alaska Sea Grant Program, Fairbanks, AK.

Hanselman, D.H., T.J. Quinn II, C. Lunsford, J. Heifetz and D.M. Clausen. 2003. Applications in adaptive cluster sampling of Gulf of Alaska rockfish. Fish. Bull. 101(3): 501-512.

Heifetz, J., D. M. Clausen, and J. N. Ianelli. 1994. Slope rockfish. *In* Stock assessment and fishery evaluation report for the 1995 Gulf of Alaska groundfish fishery. North Pacific Fishery Management Council, 605 W 4th Ave, Suite 306 Anchorage, AK 99501. pp. 5-1:5-24.

Heifetz, J., J. N. Ianelli, D. M. Clausen, D. L. Courtney, and J. T. Fujioka. 2000. Slope rockfish. In Stock assessment and fishery evaluation report for the groundfish resources of the Gulf of Alaska as projected for 2001. North Pacific Fishery Management Council, 605 W 4th Ave, Suite 306 Anchorage, AK 99501.

Heifetz, J., D.L. Courtney, D. M. Clausen, D. Hanselman, J. T. Fujioka, and J. N. Ianelli. 2002. Slope rockfish. In Stock assessment and fishery evaluation report for the groundfish resources of the Gulf of Alaska as projected for 2002. North Pacific Fishery Management Council, 605 W 4th Ave, Suite 306 Anchorage, AK 99501. pp.295-382.

Hilborn, R., Parma, A., Maunder, M. 2002. Exploitation Rate Reference Points for West Coast Rockfish: Are They Robust and Are There Better Alternatives?. North American Journal of Fisheries Management: Vol. 22, No. 1, pp. 365–375.

Ianelli, J.N. 2002. Simulation Analyses Testing the Robustness of Productivity Determinations from West Coast Pacific Ocean Perch Stock Assessment Data. North American Journal of Fisheries Management: Vol. 22, No. 1, pp. 301–310.

Ianelli, J.N. and J. Heifetz. 1995. Decision analysis of alternative harvest policies for the Gulf of Alaska Pacific ocean perch fishery. Fish. Res. 24:35-63.

Karinen, J. F., and B. L. Wing. 1987. Pacific ocean perch. In R. L. Major (editor), Condition of groundfish resources of the Gulf of Alaska region as assessed in 1986, p. 149-157. U.S. Dep. Commerce, NOAA Tech. Memo. NMFS F/NWC-119.

Leaman, B. M. 1991. Reproductive styles and life history variables relative to exploitation and management of *Sebastes* stocks. Environmental Biology of Fishes 30: 253-271.

Love, M.S., M.H. Carr, and L.J. Haldorson. 1991. The ecology of substrate-associated juveniles of the genus *Sebastes*. Environmental Biology of Fishes 30:225-243.

Love M.S, M.M. Yoklavich, and L. Thorsteinson 2002. The Rockfishes of the Northeast Pacific. University of California Press, Los Angeles.

Lunsford, C. 1999. Distribution patterns and reproductive aspects of Pacific ocean perch (*Sebastes alutus*) in the Gulf of Alaska. M.S. thesis. University of Alaska Fairbanks, Juneau Center, School of Fisheries and Ocean Sciences.

Methot, R.D. 1990. Synthesis model: An adaptable framework for analysis of diverse stock assessment data. INPFC Bull. 50: 259-289.

Quinn II, T.J., D. Hanselman, D.M. Clausen, J. Heifetz, and C. Lunsford. 1999. Adaptive cluster sampling of rockfish populations. Proceedings of the American Statistical Association 1999 Joint Statistical Meetings, Biometrics Section, 11-20.

Schnute, J.T., R. Haigh, B.A. Krishka, and P. Starr. 2001. Pacific ocean perch assessment for the west coast of Canada in 2001. Canadian research document 2001/138. 90 pp.

Seeb, L. W. and D.R. Gunderson. 1988. Genetic variation and population structure of Pacific ocean perch (*Sebastes alutus*). Can. J. Fish. Aquat. Sci. 45:78-88.

Westrheim, S.J. 1970. Survey of rockfishes, especially Pacific ocean perch, in the northeast Pacific Ocean, 1963-1966. J. Fish. Res. Bd. Canada 27: 1781-1809.

Withler, R.E., T.D. Beacham, A.D. Schulze, L.J. Richards, and K.M. Miller. 2001. Co-existing populations of Pacific ocean perch, *Sebastes alutus*, in Queen Charlotte Sound, British Columbia. Mar. Bio. 139: 1-12.

2 Applications of adaptive cluster sampling of Gulf of Alaska rockfish[4]

2.1 Introduction

In nature, populations are sometimes distributed in a patchy, rare, or aggregated manner. Conventional sampling designs such as simple random sampling (SRS) do not take advantage of this spatial information. Thompson (1990) introduced a sampling design called adaptive cluster sampling (ACS) to survey this type of distribution.

ACS, in theory, can be much more precise for a given amount of effort than conventional sampling designs (Thompson 1990). In practice, however, this is not always the case. In some cases, the variance is greatly reduced, but bias is induced from stopping rules and criterion values that are sometimes changed mid-survey (Lo et al. 1997). In 1998, I conducted a survey on Gulf of Alaska rockfish in which ACS was efficient and successful, but the gains in precision, if any, were small compared to a SRS of the same size (Quinn et al. 1999; Hanselman et al., 2001).

Recently papers about ACS have included efficiency comparisons (Christman 1996, 1997), restricted ACS (Lo et al. 1997; Brown and Manly 1998), bootstrap confidence intervals (Christman and Pontius 2000) and bias estimates (Su and Quinn 2003). However, little work has been done on determining the criterion value that, when exceeded, invokes additional sampling. In this study, I examine the details of choosing this criterion value. I illustrate this by using data from a 1999 field survey conducted for Gulf of Alaska rockfish and simulate the outcome of the experiment with different criterion values after the survey. I also compare the efficiency of ACS to SRS.

In the basic adaptive cluster sampling (ACS) design, a simple random sample (SRS) of size n is taken; if y (the variable of interest) exceeds c (a criterion value), then

[4] Hanselman, D.H., T.J. Quinn II, C. Lunsford, J. Heifetz and D.M. Clausen. 2003. Applications in adaptive cluster sampling of Gulf of Alaska rockfish. Fishery Bulletin 101:501-512. Table 2.1 was modified as requested by committee

neighborhood units are added (e.g., units above, below, left, and right in a cross pattern, Figure 2.1) to the sample. These are called network units. If any network unit has $y>c$, then its neighborhood is added. Units that do not exceed the criterion are called edge units, and sampling does not continue around them. This process continues until no units are added or until the boundary of the area is reached (Thompson and Seber 1996). Neighborhoods can be defined in any general way. The only condition is that if unit i is in the neighborhood of j, then unit j is in the neighborhood of i. The unbiasedness of the estimators relies on all neighborhood units of $y>c$ being sampled. If logistics cause the sampling to be curtailed before the sampling is complete, then biased estimators can result. For this study, all samples will be called tows because this was a trawl survey.

When little information is available to preset a fixed criterion value, order statistics are often used to choose a criterion value (Thompson and Seber 1996). The basic idea is that an initial random sample is conducted. Next, the values of the random tows are ordered, and ACS is conducted around the top ***r*** stations. The variable ***r*** is decided by the experimenter and depends on the amount of resources available and the suspected aggregation of the population. The criterion value is then set at the value of the next highest tow $(r+1)$. This was the design used in the 1998 adaptive cluster sampling survey for rockfish (Quinn et al. 1999, Hanselman et al., 2001). Using order statistics has several limitations, however. First, initial random samples must be taken before the adaptive phase can begin. This can be inefficient, because the experiment may have to move a large distance back to the previous tows that exceeded the criterion, by which time the aggregation may have moved or dispersed. In some cases, it may result in a very small criterion value that leads to an overwhelming amount of adaptive sampling around some tows. Second, achieving simple unbiased estimates of abundance is more complicated with order statistics because the criterion value is dependent on the sampling.

In this study, I address methods to avoid these limitations and illustrate these methods with a 1999 ACS survey for Gulf of Alaska rockfish. The primary target of the survey was Pacific ocean perch (*Sebastes alutus*, POP). These fish have extremely uncertain biomass estimates in the Gulf of Alaska (Heifetz et al.[5]). The estimates are based in part on a standardized stratified random survey conducted by the National Marine Fisheries Service every three years (every two since 2000). This uncertainty is likely due to their highly clustered distribution (Lunsford 1999). This has led to two independent surveys (1998, 1999) to test the benefits of ACS in sampling POP. Shortraker (*S. borealis*) and rougheye (*S. aleutianus*) rockfish combined (SR/RE) are also tested to compare the results of a population that is considered highly clustered (POP) versus one that is considered more uniformly distributed (SR/RE). SR/RE are combined because they co-occur in identical habitat and are managed as a complex.

2.2 Materials and Methods

In June 1999, ACS was carried out between 140 and 144 degrees west longitude near Yakutat in the Gulf of Alaska (Figure 2.2). Approximately 75% of sampling was directed toward the POP depth stratum (180-300m) and 25% directed toward SR/RE depths (300-450m). A 182 ft. factory trawler, the *Unimak*, was chartered to conduct trawl samples. Fishing and field operations are described in Clausen et al.[6] Duration of all trawl hauls was 15 (POP) and 30 (SR/RE) minutes on bottom. SR/RE tows were made parallel to the depth contours in a linear pattern (Figure 2.1) because the slope that SR/RE inhabit is too steep to do perpendicular tows. Travel time between all tows was recorded to examine time efficiency.

Initially, a set of systematic random tows was conducted from west to east across the entire study area to determine the criterion value. Samples were chosen systematically by longitude and distributed randomly by depth within each longitudinal

---

[5] Heifetz, J., Courtney, D.L., Clausen, D.M., Fujioka, J.T. and Ianelli, J.N. 2001. Slope rockfish. In Stock Assessment and Fishery Evaluation for the groundfish resources of the Gulf of Alaska. North Pacific Fishery Management Council, 605 W. 4th Ave, Suite 306, Anchorage, AK 99501. 72 p.

[6] Clausen, D.M., Hanselman, D.H., Lunsford, C., Quinn II, T. and Heifetz, J. 1999. Unimak Enterprise Cruise 98-01 - Rockfish adaptive sampling experiment in the central Gulf of Alaska 1998. Auke Bay Lab, NMFS, NOAA, 11305 Glacier Hwy, Auke Bay, Alaska, 99801. 49 p.

strip. This was a necessary proxy for simple random sampling, due to poorly known bathymetry in the area. If simple random latitudes and longitudes had been used, it would have resulted in sites well out of the sampling depth interval. After random sampling was completed, I compiled and examined the data to set the criterion value. Criterion values were chosen based on a hierarchy of three alternatives described below. Next, I conducted a new set of random tows from east to west across the area, in which any tows exceeding the criterion value were adaptively sampled. A distance of 0.19 km (0.1 nm) was used between all adaptive tows and the initial random tow to avoid depletion effects on the catches.

Three methods were formulated for determining a fixed criterion value *c* of POP catch-per-unit-effort (CPUE). (1) I combined and calibrated past survey and fishing data to provide the anticipated distribution of CPUE in the 1999 survey. Then I calculated the 80^{th} percentile of that distribution as the criterion value. My rationale was that this value would correspond to that obtained from order statistics. (Three networks were sampled in 1998, so the criterion value was set to the 4^{th} highest of the ordered 15 initial tows, which corresponded approximately to the 80^{th} percentile.) (2) I used the mean CPUE of past survey and fishery data, because when I compared the 80^{th} percentile criterion against the 1998 ACS survey's data, the sampling would have resulted in primarily edge units. (3) After a representative random sample was taken across the entire area in 1999, I would use the initial mean CPUE for the criterion value for the return trip. The rationale for using mean CPUE above is that in an aggregated population, the majority of the tows would be less than the mean. The actual values of the criterion chosen under each alternative are described in the results.

I chose the SR/RE criterion to be the mean CPUE of initial tows. This was assumed to be a reasonable criterion value, because if the population of SR/RE were somewhat uniform, a lower value would result in too much ACS, but mean CPUE would still be low enough to allow higher criterion values to be examined. Although I concentrated on evaluating criterion alternatives for POP, the SR/RE data are presented to illustrate that different levels of aggregation could affect how much can be gained with ACS in terms of precision and efficiency.

A major problem in applying adaptive sampling is that sampling may continue indefinitely due to a low criterion value. To limit the amount of adaptive sampling, an arbitrary stopping rule of S levels was imposed. For those strata where the cross pattern of adaptive sampling was used (POP), the stopping rule was $S = 3$ levels, allowing for a maximum of 24 adaptive tows around each high-CPUE random tow (Figure 2.1). For the strata with the linear pattern of adaptive sampling (SR/RE), the stopping rule was $S = 4$ levels, for a maximum of eight adaptive tows around each high-CPUE random tow. This differed from the previous year in which I used a stopping rule of six because I believed that the possible 30 km difference between the ends of the networks was too large for efficient sampling (Clausen[5]). In addition, no adaptive sampling extended beyond a stratum boundary. The result of adaptive sampling around each high-CPUE tow was a network of tows that extended over and, in some cases, delineated the geographic boundaries of a rockfish aggregation.

2.2.1 Statistical Methods

Statistical analysis of the results was based on adaptive cluster sampling (Thompson and Seber 1996). First, I estimated the abundance (kg/km) for the targeted rockfish species from the n initial random tows using the standard simple random sampling (SRS) estimator. Then, two adaptive estimators of abundance, a Hansen-Hurwitz estimator (HH) and a Horvitz-Thompson estimator (HT), were calculated. I computed standard error (SE) as a measure of precision. The unbiased HH estimator for the ACS mean is

$$\hat{\mu}_{HH} = \frac{1}{n}\sum_{i=1}^{n} w_i = \frac{1}{n}\sum_{i=1}^{n}\frac{y_i^*}{x_i}, \tag{1}$$

where w_i and y_i^* are the mean and total of the x_i observations in the network that intersects sample unit i, respectively. The unbiased HH estimator for the standard error is:

$$\widehat{SE}(\hat{\mu}_{HH}) = \sqrt{\frac{N-n}{Nn\,(n-1)}\sum_{i=1}^{n}\left(w_i - \hat{\mu}_{HH}\right)^2} \tag{2}$$

where N is the total number of sampling units. The HH estimator essentially replaces tows around which adaptive sampling occurred with the mean of the network of adaptive tows that exceeded the criterion CPUE.

The unbiased HT estimator for the ACS mean is

$$\hat{\mu}_{HT} = \frac{1}{N}\sum_{k=1}^{\kappa}\frac{y_k^*}{\alpha_k}, \tag{3}$$

where y_k^* is the sum of the y-values for the kth network, κ is the number of distinct networks in a sample, α_k is the probability that network k is included in the sample, and N is the total number of sampling units. If there are x_k units in the kth network, then

$$\alpha_k = 1 - \binom{N-x_k}{n}\Big/\binom{N}{n}. \tag{4}$$

where N is the total number of sampling units, n is the initial random sample and x_k is the number of units in the network. The HT estimator is based on the probability of sampling a network given the initial tows sampled and involves the number of distinct networks sampled (in contrast to the HH estimator which is based only on the initial tows). The standard error of the HT estimator is

$$SE(\hat{\mu}_{HT}) = \sqrt{\frac{1}{N^2}\left[\sum_{j=1}^{\kappa}\sum_{k=1}^{\kappa}\frac{y_j^* y_k^*}{\alpha_{jk}}\left(\frac{\alpha_{jk}}{\alpha_j \alpha_k} - 1\right)\right]} \tag{5}$$

where α_{jk} is the probability that networks j and k are both intersected by the initial point (joint selection probability) and is

$$\alpha_{jk} = 1 - \left[\binom{N-x_j}{n} + \binom{N-x_k}{n} - \binom{N-x_j-x_k}{n}\right]\Big/\binom{N}{n} \tag{6}$$

The HT estimator often outperforms other estimators as seen in simulation studies (Su and Quinn 2003). Both estimators use the network samples and initial random samples, but not the edge units. This sample size is referred to as ν' (convention established by Thompson (1990) and used in Thompson and Seber (1996)). Thompson and Seber (1996) and Salehi (1999) use the Rao-Blackwell theorem to include edge units into the estimates, which is a complex method that could theoretically result in more precise estimates. However, it had little effect for the 1998 survey data (<1% improvement, Hanselman, 2000), so these calculations are not used in this study.

When a stopping rule is used, the theoretical basis for the adaptive sampling design changes. It may result in incomplete networks that overlap and are not fixed relative to a specified criterion, changing with the pattern of the population. In contrast, the non-stopping-rule scheme has disjoint networks that form a unique partition of the population for a specified criterion. This partitioning is the theoretical basis for the unbiasedness of $\hat{\mu}_{HH}$ and $\hat{\mu}_{HT}$. Thus with a stopping rule, some bias may be introduced.

Recent simulation studies (Su and Quinn 2003) have estimated the bias induced from using a stopping rule on each estimator using order statistics, but not with a fixed criterion. Since using a fixed criterion is design unbiased, its estimate should be less biased by the stopping rule than a sample with order statistics. Therefore, I can use the Su-Quinn simulation results to approximate the maximum bias induced by the stopping rule. With a stopping rule of three and the HH estimator, the maximum positive bias is 17% for a highly aggregated simulated population. With a stopping rule of three and the HT estimator, the maximum bias is approximately 12%. Considering the design, I accepted the trade-off of relatively small bias for gains in precision and logistical efficiency.

Additionally, nonparametric bootstrap methods were adapted from Christman and Pontius (2000) using the HH version of the estimates to examine bias from this survey. Five thousand resamples were performed using n for the SRS bootstrap, and the sample size from the original criterion value of 220 kg/km (ν') was used for the ACS bootstrap. Bootstrap distributions of the data were examined for SRS and ACS designs to examine the capability of each design to demonstrate a clear central tendency.

I evaluated two hypotheses: (1) Adaptive sampling would be more effective in providing precise estimates of POP biomass than would a simple random survey design. (2) Assessment of POP abundance would benefit more from an adaptive sampling design than would SR/RE, because POP are believed to be more clustered in their distribution than SR/RE. SRS estimates were obtained from the initial random tows, and variance estimates were calculated for the initial sample size (n) and for the equivalent sample size that includes the adaptive tows but not the edge units (ν'). This makes the theoretical comparison fair, as each estimate is based on the same number of samples. Total sample size including edge units (ν) is not used in the theoretical precision comparison, but is considered when efficiency issues are examined later. These hypotheses were assessed by comparing the standard errors (SEs) of ACS to SRS. Substantial reductions in SE using ACS for POP would support the first hypothesis, while no reductions of SE using ACS for SR/RE would support the second hypothesis. This comparison is qualitative because relevant significance tests are unavailable and the two methods are different in terms of efficiency.

To evaluate different alternatives and criterion values, each network was reconstructed as if the higher criterion values had been used in the field. I also examined the trade-off between amounts of additional sampling compared to the gains in precision. A comparison was made of the SRS results using sample sizes constructed with the number of possible samples using the time/sample data I collected. In this comparison I used three new sample sizes: (1) ν_t, the number of samples that could have been done using the same amount of time for SRS if sampling time for edge units was negligible; (2) ν_e, in which the edge units had taken the same amount of time as non-edge units; (3) ν_d, in which the average distance between each tow type was used as effort instead of time with edge units included.

2.3 Results

2.3.1 Formulation of Criterion Alternatives

A total of 164 tows were conducted for the ACS experiment. Nearly all tows were made successfully, with only a few exceptions that were deemed untrawlable. These tows were conducted at the nearest trawlable bottom. I determined the POP criterion value for alternatives one and two before the survey by looking at the 1998 ACS results from a different geographic area as well as prior survey and fishery data in this study area. This was done by calculating a gear efficiency coefficient for the 1998 survey by using NMFS survey data (1993, 1996) and fishery data (1996-1998) from the observer program for the same area. This gear coefficient was then multiplied by the same data for the new area to establish the expected catches. The data used and the calculations are shown in Table 2.1. To implement alternative 3, I conducted 13 initial POP and 10 initial SR/RE random tows across the entire area. Catches from these initial tows gave the following results for each criterion alternative:

Alternative 1. The mean of the 80th percentile of the data from Table 2.1 was 641.69 kg/km. I rounded this downward to c = 540 kg/km (1000 kg/nm) for ease of operation in the field (the design was originally in kg/nm units).

Alternative 2. The mean calibrated CPUE for the area from Table 2.1 yielded a criterion value c of 220 kg/km (rounded).

Alternative 3. The mean CPUEs from the initial sample in 1999 yielded criterion values of c = 250 kg/km for POP and c = 418 kg/km for SR/RE.

The second phase of the experiment began with random tows in an east to west direction. Complete location and CPUE data for both species are located in Table 2.5. In order to analyze all alternatives, the lowest alternative was used in the field for adaptive sampling during the second phase, which resulted in the 220 kg/km criterion value for

POP from Alternative 2. For SR/RE, the criterion value was the mean CPUE of 418 kg/km from Alternative 3. The remaining alternatives were simulated following the completion of the survey.

2.3.2 POP Results

After the initial tows, 25 random tows were selected for the return trip across the area. All 25 were completed; six of those became networks of more than one unit. A total of 106 tows were completed in the POP stratum. At one of the tows that exceeded the criterion value, the captain deemed that further adaptive sampling was not feasible because of the presence of coral. Of the six networks, two overlapped forming five distinct networks. In these networks, 81 adaptive samples were taken with 49 exceeding the criterion and 32 edge units, which are not included in the estimates.

I compared the results of the original adaptive sample (Alternative 2) with the simulated results of higher criterion values (Table 2.2). The precision of simple random sample estimates using both n (number of random samples) and v' (number of random samples plus number of adaptive network samples, not edge units) was contrasted with that of the adaptive estimators described above. As the criterion value increased, n remained the same while v' and r (the number of networks) decreased. At the 220 kg/km criterion value (Alternative 2), there were substantial reductions in SE over the SRS estimators by using ACS estimators for both the n and v' sample sizes. The 250 kg/km criterion value (Alternative 3) resulted in a nearly identical sample to that of the 220 kg/km (Alternative 2) criterion value with the loss of just one network sample. Hence, the estimates were nearly identical. The HT mean estimates were slightly lower than the HH estimates for the two lowest criterion values (Alternatives 2, 3) because two networks overlapped. These networks were separated at the next higher criterion value, which aligned the estimators. The next highest criterion value of 540 kg/km (Alternative 1) showed that even though the sample size was reduced by 19 tows from the original criterion value, the ACS estimators performed nearly as well, yielding just slightly larger

SEs. When the criterion was arbitrarily doubled to 1080 kg/km, the sample size was further reduced by seven, with similar SEs to the 540 kg/km criterion value.

SRS and ACS bootstraps for POP resulted in very different distributions. Five thousand replications showed that the SRS distribution was bimodal and right skewed (Figure 2.3). The SRS mean fell on the second mode, which is more than twice the ACS mean. This bimodal distribution is driven by the presence of the very large random catch (Tow #60). If that haul is present in a bootstrap replicate, then the SRS estimate tends to be high, leading to the second mode in the bootstrap distribution. The ACS bootstrap distribution was symmetric and closely resembled a normal distribution (Figure 2.3). The average estimates of bias showed that the bias of HH was (+)4% and the bias of HT was (–)1%. The standard error had an estimated bias of (+)3% for HH and HT.

The results from this POP study and the previous 1998 study were both greatly affected by one or two very large catches, as I expected for a highly clustered population. Of interest is what happened when the largest catch was changed to a nominal catch that still exceeded the criterion value. Table 2.6 shows the results of changing haul #60 from 12,000 kg/km to 540 kg/km. When using the comparison at ν', SRS outperforms ACS in terms of SE. However, it also shows that the mean of ACS is stable because it changes little by removing a high catch, while the SRS mean is reduced by half.

2.3.3 SR/RE Results

At every third POP random tow, a tow was made in the SR/RE depth stratum. A total of 35 tows were made in the SR/RE stratum. Nine random tows yielded five distinct networks with 21 network tows and five edge units. The stopping rule was invoked for three of the five networks.

At the mean CPUE criterion (418 kg/km, Alternative 3), the adaptive estimators performed approximately the same in terms of SE compared to the SRS estimator using n (Table 2.2). Using ν', the SRS estimator yielded a lower SE than both adaptive estimators. When the criterion value increased to an arbitrarily higher value (540 kg/km) the adaptive estimators performed worse than SRS estimates using both n and ν'.

2.3.4 Time efficiency

I recorded and compared travel time between adaptive tows and simple random tows for 149 of the tows (Table 2.3). Not all the tows were used because of mechanical failure or because the factory capacity was reached. In the survey, 38 hours out of 10 days were spent in transit between sampling tows, which for a short survey was a substantial amount of the available time. For POP, substantial gains in travel-time efficiency were achieved with ACS. Average travel time for simple random tows (0.45h) was nearly triple that of adaptive tows (0.16h) for POP, which indicated that ACS can maximize sampling tows for POP when time is limited. In the SR/RE sampling, travel time for adaptive sampling (0.5h) was about the same as simple random sampling (0.49h), which was due to long linear samples that are not as close together as POP tows (Figure 2.1). Also, determination of CPUE required processing of the catch, which took various amounts of time after the completion of the tow. Due to this delay, I went to the opposite tow on the other side of the random tow when sampling SR/RE with the linear pattern, whereas there were many nearby tows when sampling POP with the cross pattern.

The travel time was added to the average tow time from gear deployment to full retrieval of 0.5 h for POP and 1.0 h for SR/RE to obtain total sampling time (per sample). Travel time was reduced by 31% using adaptive sampling (0.66h/sample) relative to simple random sampling (0.95h/sample) for POP. Sampling time efficiency for SR/RE was approximately the same for adaptive sampling (1.5h/sample) and simple random sampling (1.49h/sample) for SR/RE. These results are confounded by the fact that the random tows are spread apart because of the lesser effort applied to them. The average distance between random tows (20.2 km) was adjusted to a distance of 4.73 km as if there were 106 random tows distributed throughout the area. This distance is still larger than the average distance apart in adaptive sampling (3.22 km).

From these time and distance data, I re-estimated the precision of SRS under three new sample sizes in order to further compare the relative efficiency of ACS. I denote the sample size that could have been taken under SRS, using the same amount of time as was used during the adaptive sampling including edge units, as v_e. An alternative sample size v_t was the equivalent SRS sample size if the amount of time to sample edge units in ACS was negligible. This statistic would be useful when edge units can be determined without actually trawling them, such as hydroacoustically or visually (presence/absence). A third alternative was to find the equivalent SRS sample size v_d that would result from applying the total distance traveled in the ACS design on random stations instead. For v_e, more random POP samples would have been done than were included in the adaptive estimators (Table 2.4). The SEs of ACS were still much lower across all criterion values (Table 2.2). When I used v_t (Table 2.4), SRS was much less precise than ACS (Table 2.2). Finally, when I used distance instead of time (v_d) the results were almost exactly the same as for v_e (Table 2.4).

2.4 Discussion

My two hypotheses were that ACS would be more precise than SRS for POP and no more precise for SR/RE. The results from the 1999 field study showed that the SEs for the adaptive POP estimates were smaller than both SRS estimates using n and v', which supports the first hypothesis. One curious result is that in both 1998 and 1999, the SRS estimate of density was substantially larger than the ACS estimate, even though on average, they are both essentially unbiased. I attribute this curiosity to the more variable and skewed SRS distribution in which large sampling error on the high side is possible more often than the in the ACS estimation. Of course I fully expect that both estimates would average the same value if the experiment could be repeated many times. ACS reduced the influence of one large CPUE in the relatively small initial sample, as illustrated by the symmetric and near-normal shape of the ACS bootstrap distribution. Consequently, I conclude that ACS is a more robust estimator of density than SRS for aggregated populations. One caveat is that the precision of the estimates, if measured in terms of coefficient of variation, is similar between the two methods because of the much larger mean estimate for the SRS estimate. Monte Carlo simulations would be useful to examine the properties of the estimators under different criterion values and population densities along the lines of Su and Quinn (2003).

The SR/RE adaptive estimates all have higher SEs than the SRS estimates, supporting the second hypothesis. More than twice as many samples were directed toward POP than SR/RE, yet the POP density estimates are much more variable than for SR/RE. This much larger variability for POP is indicative of the clustering that I expected.

This experiment showed that for POP, ACS with a fixed criterion has some distinct advantages over simple random sampling and over adaptive cluster sampling with order statistics, which was used in the previous 1998 survey. Lower SEs were obtained, at one third less effort than if I just added an equivalent number of random samples. Sampling over a broader area yielded better results than the tightly stratified 1998 design. Another difference between the 1998 and 1999 design was the use of a simple random starting design and a systematic-random starting design respectively. I treated the systematic random sample in the 1999 design as if it were a simple random sample in order to obtain variance estimates. This is probably reasonable because the systematic samples were sampled randomly by depth and were far enough apart to be considered random in order. Therefore, the important difference in the results almost certainly was the stratification used in 1998, not the choice of initial sampling design (systematic sampling versus simple random sampling). This study also assumed stationary aggregations of fish. This assumption may have been better satisfied with a fixed criterion because the adaptive sampling was conducted immediately after a sample exceeded the criterion value.

While the fixed criterion eliminates bias induced by a variable criterion value, this survey still used stopping rules. If bootstrapping is a good indicator of bias, then the bias induced by stopping rules is negligible. Additionally, I have shown that a relatively high criterion value could be used to help minimize the use of these stopping rules.

This study showed that ACS is a fast and efficient way to gain a large number of samples. However, if edge units do not contribute to a better estimate and they have a similar cost or time expense as included samples, then little is gained. This deficiency points toward utilizing some method of determining edge units without actually sampling them. In fisheries surveys, this might be a double sampling design using hydroacoustics as an auxiliary variable[7] or using a design called TAPAS that hydroacoustically delineates clusters (Everson et al. 1996). In other surveys, it might be possible to detect presence of the item of interest without actually surveying the unit (such as in aerial surveys.)

An ACS design should not be attempted without some prior knowledge of the population. Populations that the design would be useful for should have an aggregated distribution which can be described by correlated variation with distance, not just a large variance relative to the mean. One way to examine the data is to fit variograms to examine spatial autocorrelation (Hanselman et al., 2001). If no prior data exist, it would not make sense to attempt ACS as an initial sampling design. I have shown that a wide range of criterion values can be used without considerable differences in the results. Therefore, only enough prior data are needed so that an adequate range of population density can be estimated. If the criterion value chosen resulted in too many or too few samples, the criterion could be adjusted, and then the design stratified into two different areas.

Most commercial fish species have survey data that can be used to determine a fixed criterion. If possible, criterion values should be determined prior to the survey, so that maximum efficiency can be attained. I have shown it may be appropriate to choose a relatively high sampling criterion such as the 80^{th} percentile of past CPUE without sacrificing estimation capabilities. This has several practical advantages. First, the design is attractive for commercial boats to perform the adaptive phase at no-cost since

---

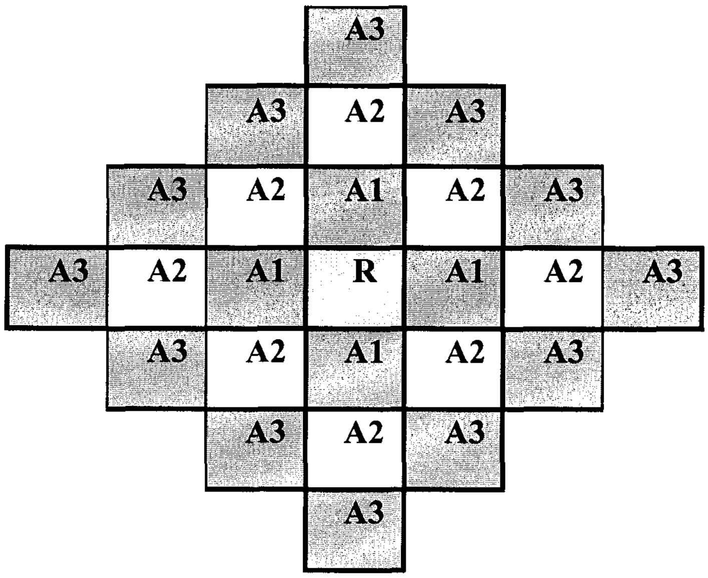
[7] Fujioka, J. 2001. Unpubl. manuscr. Using hydroacoustics and double sampling to improve rockfish abundance estimation. Auke Bay Lab, NMFS, NOAA, 11305 Glacier Hwy, Auke Bay, AK 99801, 8 p.

only large catches are sampled. The current design does not utilize the fish sampled during the survey, which in deepwater rockfish causes certain mortality. Under an adaptive design, a commercial boat would take the larger catches and could put them to use. Second, fewer overall networks would be sampled because the higher criterion would evoke less adaptive sampling, which may mean less overall sampling in the survey. Finally, precision would be gained at a minimal cost and effort. Stopping rules would be unnecessary, ensuring an unbiased estimate. However, cluster sampling is most effective when the cluster samples are as heterogeneous as possible. Therefore, caution is required to not set the criterion too high, or the resulting clusters will be either too homogenous or comprised only of edge units, leading to no improvement in the estimators. Similarly, if there are large changes in density from year to year, a fixed criterion may not be appropriate. In conclusion, adaptive cluster sampling is appropriate for surveys of highly clustered species with low temporal fluctuations, for which a fixed criterion can be determined beforehand.

2.5 Acknowledgments

I thank the crew of the F/V Unimak, in particular Captain Paul Ison and Production Manager Rob Elzig, for their excellent cooperation in this study. I also acknowledge the hard work of the scientists that participated on the cruise and NMFS personnel who prepared for the charter. I greatly appreciate the helpful comments from three anonymous reviewers that helped me refine the paper.

Cross Pattern



Linear Pattern

Figure 2.1. Maximum possible number of adaptive hauls for the cross (S = 3) and linear (S = 4) patterns with the imposition of a stopping rule. The initial random tow is denoted as "R," and the adaptive tows as "A" and their respective level number.

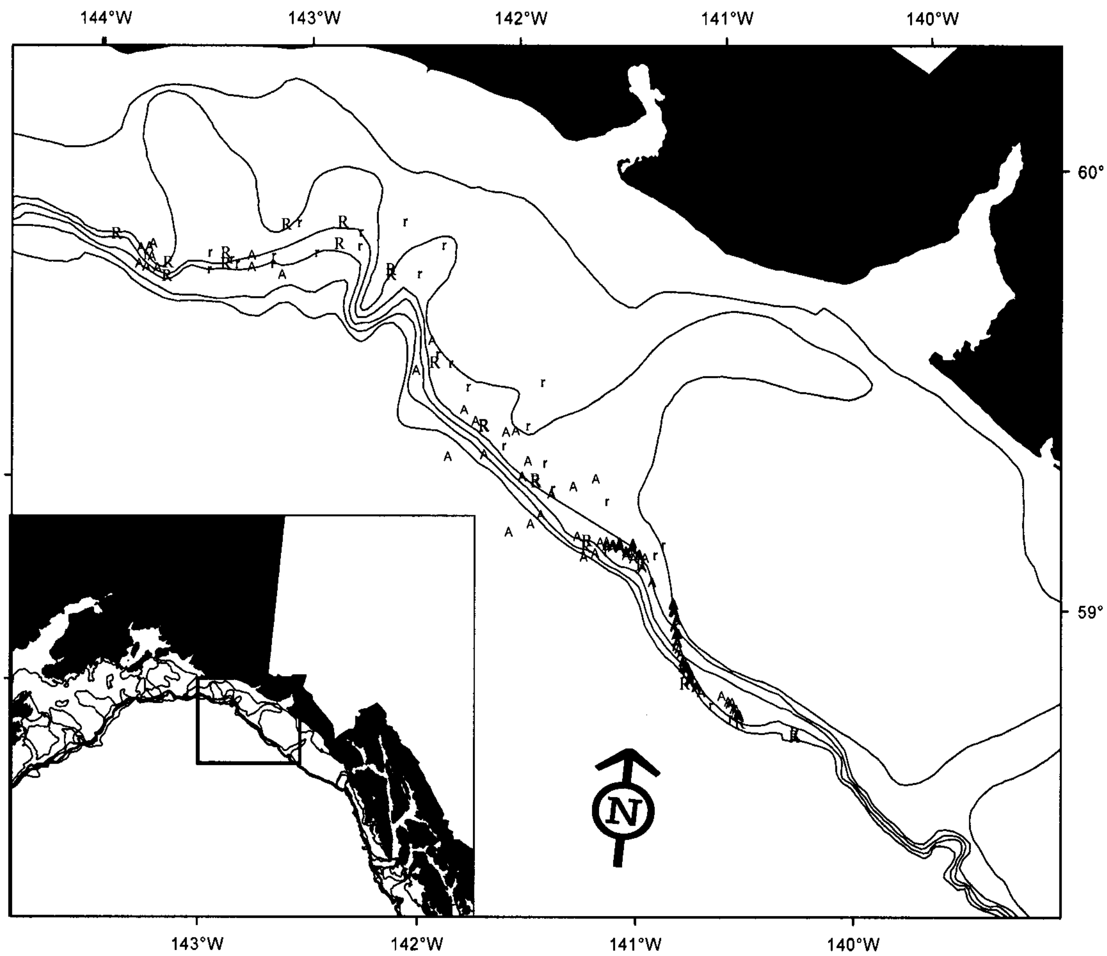


Figure 2.2. Map of sampling area in the Gulf of Alaska on the Unimak 99-01 Adaptive sampling cruise. R symbols are the initial random tows for the criterion phase, r symbols are random stations in the survey phase, A symbols are adaptive cluster samples.

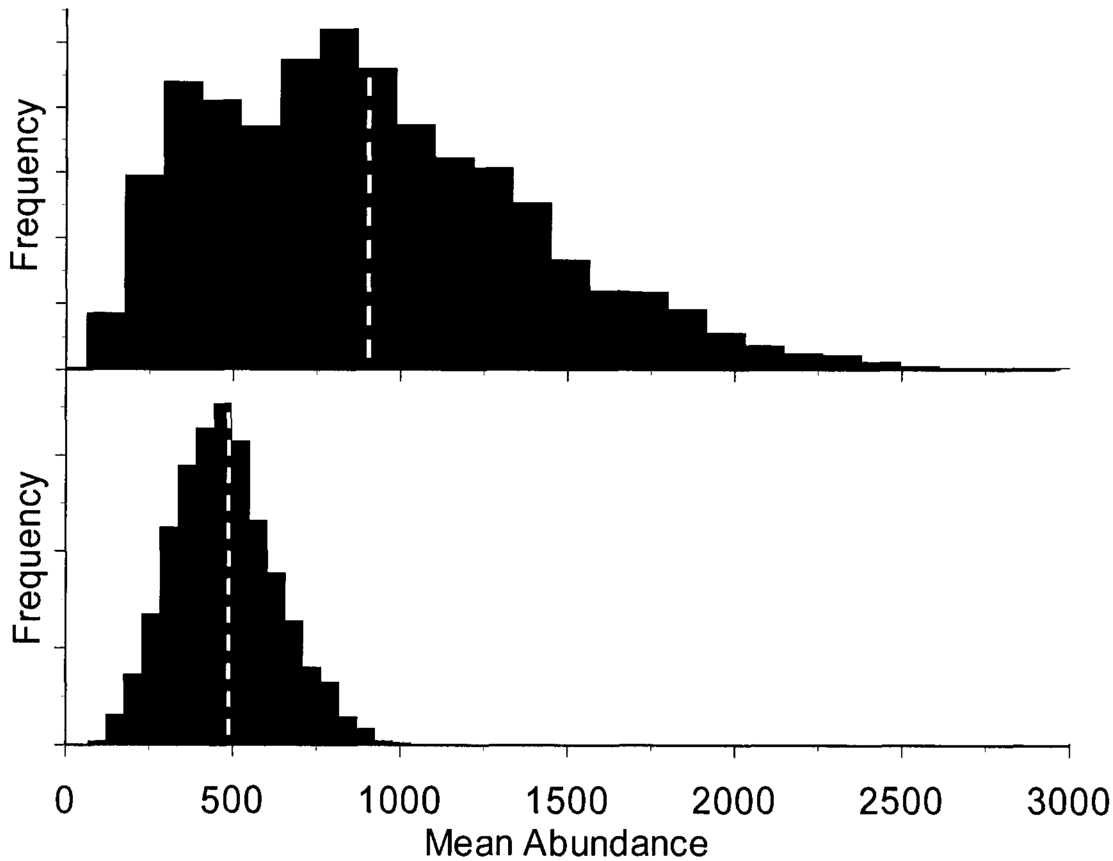


Figure 2.3. Bootstrap distributions for 1999 adaptive sampling survey (25,000 replicates). Dotted line is the mean. Top is SRS bootstrap. Bottom is ACS bootstrap (using Hansen-Hurwitz estimates).

Table 2.1. Data used to determine criterion values *c* for 1999 ACS survey. Data from a 1998 ACS survey from a different area is divided by NMFS triennial survey data and fishery data from the same area to get gear efficiencies. The mean of these gear efficiencies is then multiplied against triennial and fishery data from the new area to yield gear-calibrated CPUEs for the new area. Explained algebraically, a=ACS results from different area and year, b=Corresponding CPUES from old area, c=Gear efficiencies of Unimak, d=mean of gear efficiencies, e=Prior CPUE data from new area, f=Calibrated data for new area, g=mean of calibrated data for new area. Or: a/b=c, d=mean(c), f=d*e, g=mean(f).

Data Source	Year	Mean CPUE (kg/km)	80%	*n*
ACS Results from different area and year	1998	284.94	223.92	57
(Divided by)			÷	
Corresponding	Triennial 1993	38.36	7.89	50
previous area	Triennial 1996	46.64	27.33	51
CPUEs from triennial and fishery data	Fishery 1996-98	30.64	14.03	434
(Equals)			=	
	Triennial 1993	7.44	28.18	
Gear efficiency of	Triennial 1996	6.12	8.14	
Unimak	Fishery 1996-98	9.32	15.85	
Average gear efficiency	**Mean**	**7.63**	**17.39**	
(Multiplied by)			x	
Prior CPUE data	Triennial 1993	40.32	46.74	29
from area for 1999	Triennial 1996	26.50	33.50	25
ACS survey	Fishery 1996-98	19.61	30.47	190
			=	
(Equals)	Triennial 1993	307.52	812.67	29
Calibrated CPUE	Triennial 1996	202.06	582.52	25
data for 1999 ACS survey	Fishery 1996-98	149.57	529.90	137
Criterion value *c*	**Mean**	**219.71**	**641.69**	

Table 2.2. Summary of rockfish density estimates ($\hat{\mu}$) and standard errors (SE) for 1999 adaptive cluster sampling experiment. *c* is criterion value, *r* is number of adaptive networks, *n* is initial sample size, v' is adaptive sampling size (excluding edge units). SRS is simple random sampling estimator, HH is Hansen-Hurwitz adaptive estimator and HT is Horvitz-Thompson adaptive estimator.

	S. alutus				*S. borealis/aleutianus*	
Alternative	2	3	1	-----	3	-----
c (kg/km)	>220	>250	>540	>1080	>418	>540
r	6	6	5	3	5	3
n	25	25	25	25	9	9
v'	74	73	55	48	30	14
$\hat{\mu}_{SRS}$	904	904	904	904	447	447
SE_n	496	496	496	496	115	115
$SE_{v'}$	288	290	334	358	63	92
$\hat{\mu}_{HH}$	498	501	566	526	511	486
SE	166	167	192	197	128	141
$\hat{\mu}_{HT}$	471	472	567	527	511	486
SE	167	167	192	197	128	141

Table 2.3. Comparisons of time/haul and time/sample of adaptive sampling against simple random sampling for Pacific ocean perch and shortraker/rougheye rockfish on a 1999 adaptive sampling cruise. Time/travel is travel time between tows in hours; time/sample is travel time plus haul time in hours. Distance between is average travel distance (km) between two adaptive stations and between two random stations. Adjusted distance reflects what the distance would be if the random sample size was increased to 106.

	S. alutus		*S. borealis/aleutianus*	
	Random	Adaptive	Random	Adaptive
Time (h)	10.4	11.4	4.4	12.0
Hauls	23	72	9	24
Time/travel	0.45	0.16	0.49	0.50
Time/sample	0.95	0.66	1.49	1.50
Distance between	20.2	3.22		
Adjusted distance	4.73	3.22		

Table 2.4. Comparison of simple random sampling (SRS) precision estimates with the inclusion of time and distance information. c is the criterion value. ν' is the original adaptive cluster sampling adjusted sample size. ν_e is the time-adjusted sample size, including edge units. ν_t is the time-adjusted sample size with edge unit cost set to zero. ν_d is the distance-adjusted sample size including edge units. $\hat{\mu}$ is the mean SRS density estimate, SE is the standard error for that sample size.

	Criterion value c (kg/km)			
Parameter	>220	>250	>540	>1080
$\hat{\mu}$	904	904	904	904
ν'	74	73	55	48
SE	294	296	341	365
ν_e	81	80	67	55
SE	281	283	309	341
ν_t	59	58	46	41
SE	329	332	373	395
ν_d	80	79	67	54
SE	283	285	309	344

Table 2.5. CPUE (kg/km) data from 1999 adaptive cluster sampling survey. CPUE is in kg/km. The format of "Adaptive 26-1" corresponds to first adaptive tow around haul #26.

Summary table

Tow type	Initial Random	2^{nd} Phase Random	Adaptive Network	Adaptive Edge Unit	Total*
POP	13	25	49	32	106 (119)
SR/RE	10	9	21	5	35 (45)
Total	23	34	70	37	141 (164)

* Values in parenthesis include initial random tows that are not included in estimation results. POP is *Sebastes alutus*, SR/RE are *Sebastes borealis/aleutianus*.

Table 2.5 (continued).

Tow	Latitude	Longitude	Tow type	POP CPUE	SR/RE CPUE
			Criterion-determining random tows		
3	59.59	-143.81	POP Random	39.3	43.7
4	59.54	-143.55	POP Random	49.2	13.7
5	59.51	-143.55	SR/RE Random	3.4	870.9
6	59.58	-143.28	POP Random	174.8	112.0
7	59.56	-143.28	SR/RE Random	17.7	582.3
8	59.67	-143.01	POP Random	72.7	21.0
9	59.69	-142.75	POP Random	21.3	6.1
10	59.64	-142.75	SR/RE Random	6.3	6.3
11	59.60	-142.49	POP Random	9.6	36.2
12	59.59	-142.48	SR/RE Random	3.8	608.0
13	59.40	-142.22	POP Random	20.7	113.0
14	59.28	-141.96	POP Random	25.3	394.4
15	59.27	-141.96	SR/RE Random	19.1	713.1
16	59.17	-141.68	POP Random	185.4	68.5
17	59.16	-141.68	SR/RE Random	24.9	48.5
18	59.04	-141.41	SR/RE Random	1.7	450.4
19	59.03	-141.41	POP Random	196.5	21.9
20	59.01	-141.14	SR/RE Random	30.0	676.9
21	58.78	-140.88	POP Random	2271.6	0.0
22	58.75	-140.88	SR/RE Random	65.9	80.6
23	58.67	-140.61	POP Random	80.6	101.1
24	58.66	-140.35	POP Random	98.2	55.0
25	58.66	-140.35	SR/RE Random	21.2	140.5
			Begin Adaptive Random Tows		
26	58.70	-140.64	POP Random	576.7	0.0
27	58.68	-140.65	SR/RE Random	16.3	115.8

Tow	Latitude	Longitude	Tow type	POP CPUE	SR/RE CPUE
28	58.73	-140.71	POP Adapt. 26-1	138.1	12.0
29	58.72	-140.65	POP Adapt. 26-2	138.4	9.7
30	58.69	-140.62	POP Adapt. 26-3	2294.2	0.0
31	58.70	-140.64	POP Adapt. 26-4	290.1	0.4
32	58.70	-140.63	POP Adapt. 26-8	334.8	0.0
33	58.69	-140.62	POP Adapt. 26-9	56.5	21.2
34	58.69	-140.63	POP Adapt. 26-10	16.4	1.9
35	58.71	-140.67	POP Adapt. 26-11	20.7	3.7
36	58.72	-140.67	POP Adapt. 26-12	30.2	1.0
37	58.69	-140.61	POP Adapt. 26-18	1299.4	1.2
38	58.69	-140.61	POP Adapt. 26-17	965.0	55.9
39	58.70	-140.75	POP Random	62.0	148.0
40	58.76	-140.85	POP Random	3591.0	58.4
41	58.79	-140.89	POP Adapt. 40-1	5934.1	0.0
42	58.77	-140.86	POP Adapt. 40-2	4521.0	0.0
43	58.74	-140.83	POP Adapt. 40-3	515.7	9.1
44	58.76	-140.86	POP Adapt. 40-4	4453.7	37.3
45	58.79	-140.90	POP Adapt. 40-5	1338.8	0.0
46	58.79	-140.88	POP Adapt. 40-6	393.9	0.0
47	58.77	-140.86	POP Adapt. 40-7	109.4	0.0
48	58.75	-140.82	POP Adapt. 40-8	85.0	0.0
49	58.73	-140.80	POP Adapt. 40-9	67.9	0.1
50	58.74	-140.83	POP Adapt. 40-10	128.0	17.6
51	58.76	-140.86	POP Adapt. 40-11	1597.3	0.0
52	58.78	-140.89	POP Adapt. 40-12	268.5	3.8
53	58.80	-140.90	POP Adapt. 40-24	1282.9	0.0
54	58.81	-140.92	POP Adapt. 40-13	2304.4	0.0
55	58.80	-140.90	POP Adapt. 40-14	776.2	0.0
56	58.79	-140.88	POP Adapt. 40-15	882.6	0.0
57	58.75	-140.86	POP Adapt. 40-22	168.1	2.7
58	58.78	-140.89	POP Adapt. 40-23	253.9	0.2
59	58.83	-140.95	SR/RE Random	24.1	290.2
60	58.88	-140.95	POP Random	12001.5	0.0
61	58.87	-140.96	POP Adapt. 60-4	10659.3	0.0
62	58.91	-140.97	POP Adapt. 60-1	1179.0	0.0
63	58.89	-140.95	POP Adapt. 60-2	3050.4	0.0
64	58.86	-140.95	POP Adapt. 60-3	2984.7	0.0

Tow	Latitude	Longitude	Tow type	POP CPUE	SR/RE CPUE
65	58.86	-140.95	POP Adapt. 60-10	3590.4	0.0
66	58.88	-140.96	POP Adapt. 60-11	1086.9	0.0
67	58.91	-140.98	POP Adapt. 60-12	1311.7	8.7
68	58.92	-140.98	POP Adapt. 60-5	1581.0	0.0
69	58.91	-140.96	POP Adapt. 60-6	4148.4	0.0
70	58.89	-140.95	POP Adapt. 60-7	1297.4	0.0
71	58.86	-140.94	POP Adapt. 60-8	214.1	0.0
72	58.84	-140.94	POP Adapt. 60-9	2190.3	0.0
73	58.84	-140.94	POP Adapt. 60-20	1502.2	0.0
74	58.83	-140.93	POP Adapt. 60-19	2828.9	0.0
75	58.84	-140.93	POP Adapt. 60-18	102.9	0.0
76	58.86	-140.94	POP Adapt. 60-17	46.6	0.0
77	58.89	-140.95	POP Adapt. 60-16	27.8	0.0
78	58.89	-140.95	POP Adapt. 60-15	53.4	0.0
79	58.92	-140.97	POP Adapt. 60-14	495.7	0.0
80	58.93	-140.98	POP Adapt. 60-13	1323.4	0.0
81	59.05	-141.05	POP Random	1448.8	0.4
82			Coral Hangup	N/A	N/A
83	59.03	-141.08	POP Random	560.6	102.8
84	59.03	-141.19	POP Random	283.6	298.5
85	59.04	-141.19	POP Adapt. 83-1	1119.7	101.3
86	59.04	-141.26	POP Adapt. 83-2	1407.0	21.7
87	59.02	-141.22	POP Adapt. 83-3	398.1	29.2
88	59.03	-141.16	POP Adapt. 83-4	264.6	87.0
89	59.05	-141.20	POP Adapt. 83-5	416.6	47.3
90	59.04	-141.29	POP Adapt. 83-6	2186.1	7.0
91	59.04	-141.25	POP Adapt. 83-7	482.0	8.7
92	59.03	-141.22	POP Adapt. 83-8	115.2	36.6
93	59.02	-141.19	POP Adapt. 83-9	182.5	36.4
94	59.02	-141.13	POP Adapt. 83-10	41.4	45.5
95	59.02	-141.16	POP Adapt. 83-11	29.2	41.1
96	59.04	-141.20	POP Adapt. 83-12	261.4	80.6
97	59.04	-141.25	POP Adapt. 83-24	109.3	32.0
98	59.04	-141.29	POP Adapt. 83-23	62.0	69.4
99	59.05	-141.26	POP Adapt. 83-13	186.4	56.2
100	59.05	-141.32	POP Adapt. 83-14	443.8	4.5
101	59.04	-141.29	POP Adapt. 83-15	1497.1	5.4

Tow	Latitude	Longitude	Tow type	POP CPUE	SR/RE CPUE
102	59.04	-141.25	POP Adapt. 83-16	892.0	21.4
103	59.03	-141.22	POP Adapt. 83-17	604.8	26.1
104	59.03	-141.16	POP Adapt. 84-3	123.5	91.4
105	59.03	-141.22	POP Adapt. 84-4	129.3	285.3
106	59.04	-141.26	POP Adapt. 84-1	231.2	602.5
107	59.02	-141.32	SR/RE Random	49.3	721.9
108	59.05	-141.26	POP Adapt. 84-5	214.6	1408.9
109	59.04	-141.35	POP Adapt. 84-6	215.0	123.6
110	59.04	-141.31	POP Adapt. 84-12	61.5	664.5
111	59.04	-141.32	SR/RE Adapt. 107-1	57.5	758.1
112	59.02	-141.37	SR/RE Adapt. 107-2	0.0	490.7
113	59.05	-141.20	SR/RE Adapt. 107-3	0.0	408.6
114	59.01	-141.42	SR/RE Adapt. 107-4	0.0	669.1
115	59.00	-141.14	SR/RE Adapt. 107-6	0.0	760.8
116	58.97	-141.09	SR/RE Adapt. 107-8	0.0	1540.6
117	58.11	-141.06	SR/RE Random	0.0	443.2
118	59.14	-141.60	SR/RE Adapt. 117-1	0.0	1052.8
119	59.09	-141.64	SR/RE Adapt. 117-2	0.0	1042.0
120	59.16	-141.50	SR/RE Adapt. 117-3	51.3	621.6
121	59.07	-141.69	SR/RE Adapt. 117-4	25.7	2096.7
122	59.05	-141.46	SR/RE Adapt. 117-6	68.4	480.5
123	59.19	-141.40	SR/RE Adapt. 117-5	41.2	924.3
124	59.21	-141.73	SR/RE Adapt. 117-7	189.0	731.9
125	59.04	-141.78	SR/RE Adapt. 117-8	82.3	772.2
126	59.14	-141.34	POP Random	61.9	4.8
127	59.15	-141.60	POP Random	82.6	55.8
128	59.21	-141.65	POP Random	68.5	8.1
129	59.29	-141.75	POP Random	84.6	0.0
130	59.23	-141.85	SR/RE Random	6.1	1024.1
131	59.27	-141.85	SR/RE Adapt. 130-1	2.6	626.9
132	59.21	-141.94	SR/RE Adapt. 130-2	1.5	451.9
133	59.27	-141.81	SR/RE Adapt. 130-3	4.2	2208.3
134	59.28	-142.00	SR/RE Adapt. 130-5	7.4	1605.6
135	59.31	-142.06	SR/RE Adapt. 130-7	5.0	1305.2
136	59.19	-142.11	SR/RE Adapt. 130-4	0.0	432.4
137	59.17	-141.75	SR/RE Adapt. 130-6	1.6	457.4
138	59.39	-141.70	POP Random	181.8	25.9

Tow	Latitude	Longitude	Tow type	POP CPUE	SR/RE CPUE
139	59.36	-142.05	POP Random	62.9	12.2
140	59.40	-142.15	SR/RE Random	3.7	772.3
141	59.45	-142.25	SRRE Adapt. 140-1	1.1	222.7
142	59.38	-142.31	SRRE Adapt. 140-2	0.0	209.0
143	59.42	-142.22	POP Random	177.2	36.0
144	59.67	-142.25	POP Random	45.4	33.5
145	59.60	-142.35	POP Random	8.3	117.8
146	59.71	-142.45	POP Random	4.3	32.0
147	59.67	-142.65	SR/RE Random	2.0	47.0
148	59.64	-142.65	POP Random	18.0	50.8
149	59.67	-142.95	POP Random	34.2	3.4
150	59.61	-142.85	POP Random	125.0	18.8
151	59.57	-143.05	SR/RE Random	3.6	530.5
152	59.59	-143.05	POP Random	139.0	39.7
153	59.56	-143.15	SR/RE Adapt. 151-1	5.1	555.2
154	59.59	-143.16	SR/RE Adapt. 151-2	2.6	255.5
155	59.55	-143.00	SR/RE Adapt. 151-3	0.0	314.5
156	59.56	-143.22	POP Random	23.5	567.4
157	59.57	-143.25	POP Random	43.3	399.3
158	59.54	-143.35	SR/RE Random	9.3	82.2
159	59.58	-143.36	POP Random	74.9	493.0
160	59.55	-143.45	POP Random	2838.5	1.8
161	59.57	-143.65	POP Adapt. 160-1	1674.5	54.5
162	59.53	-143.69	POP Adapt. 160-2	2912.8	1.8
163	59.55	-143.63	POP Adapt. 160-3	196.5	0.0
164	59.52	-143.65	POP Adapt. 160-4	148.2	0.5
165	59.52	-143.60	POP Adapt. 160-5	75.6	21.0
166	59.58	-143.63	POP Adapt. 160-6	863.1	9.4
167	59.56	-143.69	POP Adapt. 160-7	41.3	0.0

Table 2.6. Results of estimation with Haul #60 changed from 12000 kg/km to 540 kg/km. *c* is criterion value (kg/km), $\hat{\mu}$ is mean Pacific ocean perch density (kg/km) for each estimator, *n* is random sample size, ν is adaptive sample size without edge units. SE is standard error of the mean.

c (kg/km)	>220	>250	>540	>1080
$\hat{\mu}_{srs}(n)$	445	445	445	445
SE	179	179	179	179
SE (ν)	104	104	104	104
$\hat{\mu}_{HH}$	470	473	535	412
SE	148	149	175	158
$\hat{\mu}_{HT}$	442	443	536	413
SE	149	149	175	158

2.6 References

Brown, J.A. and B.J.F. Manly. 1998. Restricted adaptive cluster sampling. Environ. Ecol. Stat. 5:49-63.

Christman, M.C. 1996. Comparison of efficiency of adaptive sampling in some spatially clustered populations. American Statistical Association 1996 Proceedings of the Section on Statistics and the Environment, 122-126.

Christman, M.C. 1997. Efficiency of some sampling designs for spatially clustered populations. Environmetrics 8:145-166.

Christman, M.C. and J.S. Pontius. 2000. Bootstrap confidence intervals for adaptive cluster sampling. Biometrics 56:503-510.

Everson, I., M. Bravington and C. Goss. 1996. A combined acoustic and trawl survey for efficiently estimating fish abundance. Fish. Res. 26:75-91.

Hanselman, D.H., T.J. Quinn II, C. Lunsford, J. Heifetz and D.M. Clausen. 2001. Spatial inferences of adaptive cluster sampling on Gulf of Alaska rockfish. Proceedings of the 17th Lowell-Wakefield Symposium: Spatial Processes and Management of Marine Populations, 303-325.

Lo, N., D. Griffith and J.R. Hunter. 1997. Using a restricted adaptive cluster sampling to estimate Pacific hake larval abundance. Calif. Coop. Oceanic Fish. Invest. Rep. 38:103-113.

Lunsford, C. 1999. Distribution patterns and reproductive aspects of Pacific ocean perch (*Sebastes alutus*) in the Gulf of Alaska. M.S. thesis. Univ. of Alaska Fairbanks, Juneau Center, School of Fisheries and Ocean Sciences, 154 p.

Quinn II, T.J., D.H. Hanselman, D.M. Clausen, J. Heifetz and C. Lunsford. 1999. Adaptive cluster sampling of rockfish populations. Proceedings of the American Statistical Association 1999 Joint Statistical Meetings, Biometrics Section, 11-20.

Salehi, M.M. 1999. Rao-Blackwell versions of the Horvitz-Thompson and Hansen-Hurwitz in adaptive cluster sampling. Environ. Ecol. Stat. 6:83-195.

Su, Z. and T.J. Quinn II. 2003. Estimator bias and efficiency for adaptive cluster sampling with order statistics and a stopping rule. Environ. Ecol. Stat. 10: 17-41.

Thompson, S.K. 1990. Adaptive cluster sampling. J. Am. Stat. Assoc. 412:1050-1059.

Thompson, S.K. and Seber, G.A.F. 1996. Adaptive Sampling. Wiley, New York, 265 p.

3 Simulations in adaptive cluster sampling[8]

3.1 Introduction

Adaptive cluster sampling (ACS) has recently been the focus of much attention (see Thompson and Seber 1996 and Environmental and Ecological Statistics vol. 10 (2003)). The main attraction of this sampling design is its ability to gain survey precision for highly aggregated populations with less effort than surveying additional random stations. If the sampling distribution is highly skewed, then the adaptive estimators are more likely to exhibit central tendency than conventional random sampling because the distribution of sample means is less skewed (Conners and Schwager 2002, Hanselman et al. 2003).

Rockfish in Alaska are typically surveyed with trawl gear. A basic ACS survey for rockfish starts with any type of random sampling design, such as simple or stratified random sampling, until a random trawl sample, or tow, surpasses a criterion value (c). Next, samples are conducted in the neighborhood of that random tow until the catch of a tow drops below the criterion. The samples that do not exceed the criterion on the periphery of the network are called edge units. These values are not used in the estimates. The resultant network can be a variety of different shapes, including the cross and linear patterns (Figure 3.1). In part one, I use the cross pattern (left, right, top, bottom). These shapes can determine the efficiency of the ACS design relative to random sampling. In Brown (2003), the cross pattern has a higher maximum efficiency than the linear design, which has the smallest maximum efficiency gain over simple random sampling, but also is less likely to be much worse than simple random sampling if the population sampled is not rare or clustered.

---

[8] Part 1 of 2 of: Hanselman, D.H. and Quinn, T.J. II. In Press Sampling Rockfish Populations: Adaptive Sampling and Hydroacoustics. *In* Sampling Rare or Elusive Populations: Challenges and Choices edited by William Thompson. Island Press. Split up by request of committee.

Most of the literature on ACS concerns the theory and simulation of the design. Few have field tested it on marine populations (Thompson and Seber 1996). I conducted two ACS field experiments on rockfish in the Gulf of Alaska. The first experiment (Hanselman et al. 2001) used a stratified-random design with a criterion value determined by order statistics. Order statistics can be a useful way of setting the criterion value when little is known *a priori* about the population being sampled. Basically, an initial random sample is conducted, and then the catch-per-unit-effort (CPUE, kg/km^2) values are ordered from highest to lowest. The scientist chooses how many of the top stations to adaptively sample and uses the next highest CPUE value as the criterion. I chose a relatively small area with four strata determined roughly by habitat type. A small initial sample size of u~15 was conducted in each stratum, and then the criterion value was set by ordering the random tows by CPUE values. I adaptively sampled the top three stations using the fourth highest station's CPUE as the criterion value. An example of the results from one stratum is below:

	Adaptive Stations			Criterion	Remainder of random initial random sample										
Order	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
CPUE	5951	4681	3888	464	332	311	194	125	108	100	83	54	51	28	27

The standard errors of the resulting abundance estimates from the 1998 study were no better than if I had simply added more random samples (Figure 3.2a).

The second experiment (Hanselman et al. 2003) used a systematic-random sample over a broad area in the Gulf of Alaska. In this experiment I used a fixed criterion that was determined from prior survey data. This technique allowed adaptive sampling to begin immediately when the criterion was exceeded. This method increased logistical efficiency because I did not have to finish the initial random sample before adaptively sampling the station. By beginning the adaptive sampling immediately after the random tow, the assumption of geostationarity was better satisfied, because fish movement was less likely during the sampling time frame. When I used this type of criterion value and a larger unstratified area, I obtained slightly higher precision for the abundance estimates than if I had just added more random samples (Figure 3.2b). Another important result of the second study was that a relatively high criterion value could be used. This resulted in substantial savings in effort with only a small trade-off in precision.

The main result from the two rockfish studies were that most of the gains achieved by ACS were in terms of logistical efficiency, and the gains expected in statistical efficiency were minimal. This led to exploring the properties of ACS using simulated data from another author and another data set representing actual survey data. I used simulations to investigate why ACS did not work well for *S. alutus* during previous studies and when it might work well.

3.2 Materials and Methods

Two populations were simulated for analysis. Each population was modeled using a 40 x 40 grid (U=1600). The first population type was modeled after the "highly-aggregated" population used in Su and Quinn (2003) and the second type was modeled to reflect the characteristics of a population of *S. alutus*. Both were generated by a Poisson cluster process (Diggle 2001). The characteristics of the two populations are shown in Table 3.1. The populations are compared by their overall coefficient of variation (CV) and their proportion of zero (Pzr) cells in the grid. I set the means equal to the mean used in Su and Quinn (2003). The CV and Pzr for population 1 were set equal to those used Su and Quinn (2003) and those for population 2 were set equal to values from survey data for *S. alutus* from the 2001 NMFS biennial survey.

Each population was then sampled with six initial sample sizes (u_1) ranging between 40 and 240, representing sampling fractions between 2.5% and 15% for the 1600 blocks in the population. I performed one thousand replications for each different initial sample size. I set the criterion value for network sampling at the population mean (μ) and one-half the population mean (0.5μ). Adaptive sampling was performed in the cross pattern. I then calculated summary statistics for simple random sampling (SRS) and adaptive cluster sampling (ACS) (the Hansen-Hurwitz (HH) and Horvitz-Thompson (HT) estimators). Details of these estimators are reviewed in Section 2.2.1 of this thesis.

The efficiency of ACS was then examined by comparing the results from the two different populations. For comparison, the relative efficiency was computed as the variance of SRS divided by the variance of ACS; a value above one indicated a more efficient ACS design, whereas a value below one was less efficient. The designs were compared using two different final sample sizes: 1) ν is the final sample size including edge units, which is the practical level when all units must be surveyed 2) ν' is the final sample size without edge units that compares the designs at the same theoretical level (the sample size used in the estimators). I also investigated the effect of initial sample size on both efficiency and final sample sizes.

3.3 Results

Results of the adaptive sampling simulations were sensitive to both initial sample size and criterion value. The relative efficiency of ACS compared to SRS varied depending on which final sample size (ν or ν') was used. In population 1 (Figure 3.3) at ν (final sample size including edge units), the HT estimator was more efficient than SRS only at large initial sample sizes (>160) and criterion μ. At ν' (final sample size without edge units) for population 1, ACS was slightly more efficient than SRS regardless of initial sample size (Figure 3.3), except in the HH estimator at criterion μ. In population 2, the HT estimator was more efficient than SRS at an initial sample size between 120 and 160, depending on the criterion value used, whereas the HH estimator was never more efficient than SRS (Figure 3.4). The HT estimator was more efficient than the HH estimator, all factors held equal. The HH estimator rarely attained efficiency greater than SRS (Figures 3.3, 3.4).

Results for the HT estimator were sensitive to the criterion value chosen in both populations. For population 1, the choice of mean CPUE for the criterion value usually yielded a more efficient design at any of the initial sample sizes. The difference in final sample size between the two criterion values was much greater for population 1 (Figure

3.3). For population 2, the criterion value of one-half mean CPUE was generally a better choice in terms of maximum possible efficiency, but required a very large initial sample size. For both populations the criterion value of mean CPUE was as efficient as the one-half CPUE when the same final sample size was the same (Figures 3.4, 3.5). For example, in Figure 3.4, the initial sample size of 200 for criterion μ, had approximately the same final sample size and efficiency as an initial sample size of 160 for criterion 0.5μ.

3.4 Discussion

This study illustrates simulations of situations where adaptive cluster sampling (ACS) does not perform well on large marine populations. The results for the strongly aggregated population showed that small improvements in precision can be gained at most initial sample sizes, with either criterion value. This population was less variable than the *S. alutus*-like population with more zero catches. ACS for the *S. alutus*-like population was less efficient than SRS at lower sample sizes at either criterion value but became much more efficient at a high initial sampling fraction (>10% of population). This result is consistent with the two field experiments performed on *S. alutus,* where ACS provided relatively small gains in precision (Hanselman et al. 2001, 2003) when small initial sampling fractions were used. The simulations also confirmed results determined in Hanselman et al. (2003) that a higher criterion value reduces sampling effort with small losses in precision. Su and Quinn (2003) also show large possible gains with the HT estimator on a smaller population (U=400) when greater than 10% of the sampling frame is included in the initial sample. Therefore, to reap the potential large gains in precision with an adaptive design, a trawl survey would need to cover 10% or more of the sampling units for the population for the initial sample alone. Unfortunately, in an area as large as the Gulf of Alaska, the current trawl survey covers <0.5% of the possible sampling units. Because this survey is already large in terms of resources (>204 vessel days, >1000 scientist days), a large increase in coverage is unlikely.

These studies on the utility of adaptive cluster sampling for large clustered marine populations have shown that the design may not be appropriate when a much larger initial sample size is unlikely to be obtained. The design might be to obtain a good estimate of a small area where a large initial sampling fraction is possible. I have also shown that adaptive cluster sampling can capture fine-scale variability when insufficient information is available for stratification to capture large-scale variability (Hanselman et al. 2001).

3.4.1 Conclusions and future work

My studies on applying adaptive cluster sampling to *S. alutus* have shown that it is mainly good for two applications. If a population estimate is needed for a small area like a marine reserve, then adaptive cluster sampling would be a good choice. Additionally, if the goal is to maximize the number of samples that can be taken in a survey in order to gain biological information and to delineate clusters of fish, then adaptive cluster sampling would work well. However, for a large-scale marine population like the Gulf of Alaska, a stratified random design that is optimal for *S. alutus* or rockfish in general would probably work best.

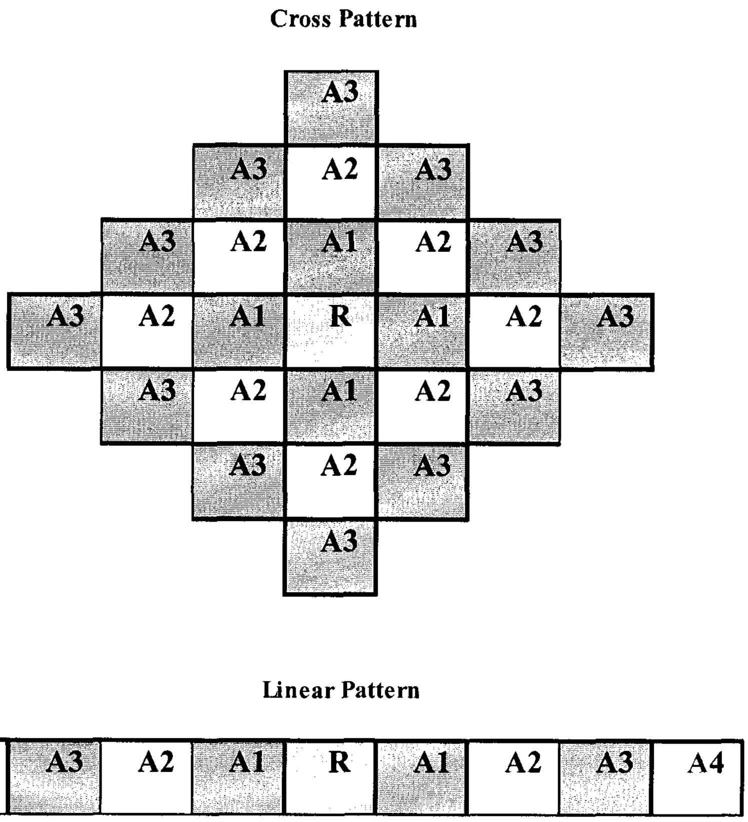


Figure 3.1. Two example network patterns for adaptive cluster sampling. R is the initial random station that exceeds the criterion value. A1 is the first level of adaptive sampling.

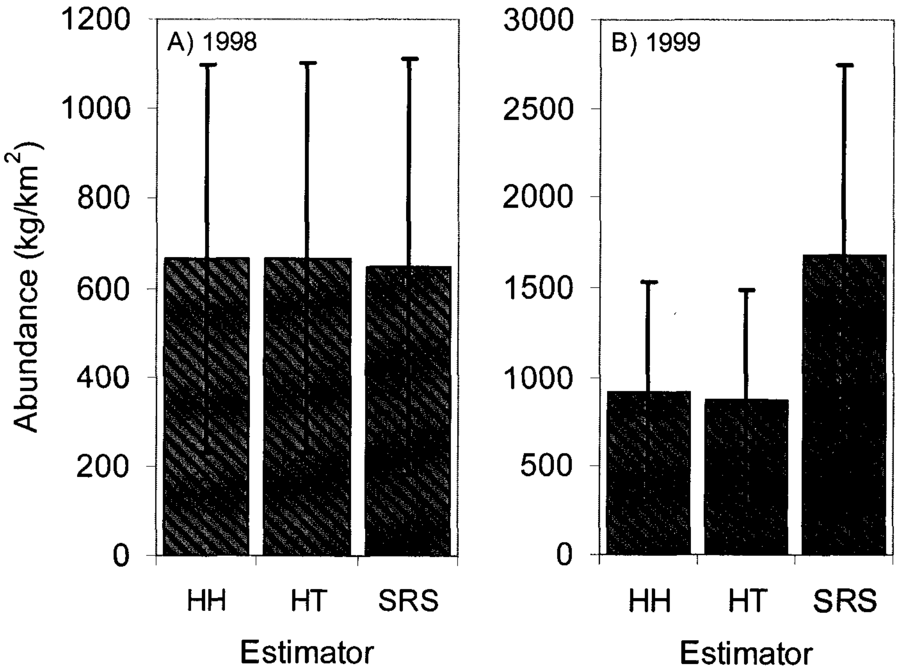


Figure 3.2. Results of 1998 (a) and 1999 (b) adaptive cluster sampling experiments for *S. alutus*. HH=Hansen-Hurwitz estimator, HT=Horvitz-Thompson estimator, SRS=Simple random sampling estimator. Sample size of ν (initial sample size plus adaptive sample minus edge units) is used for comparison at the same theoretical level.

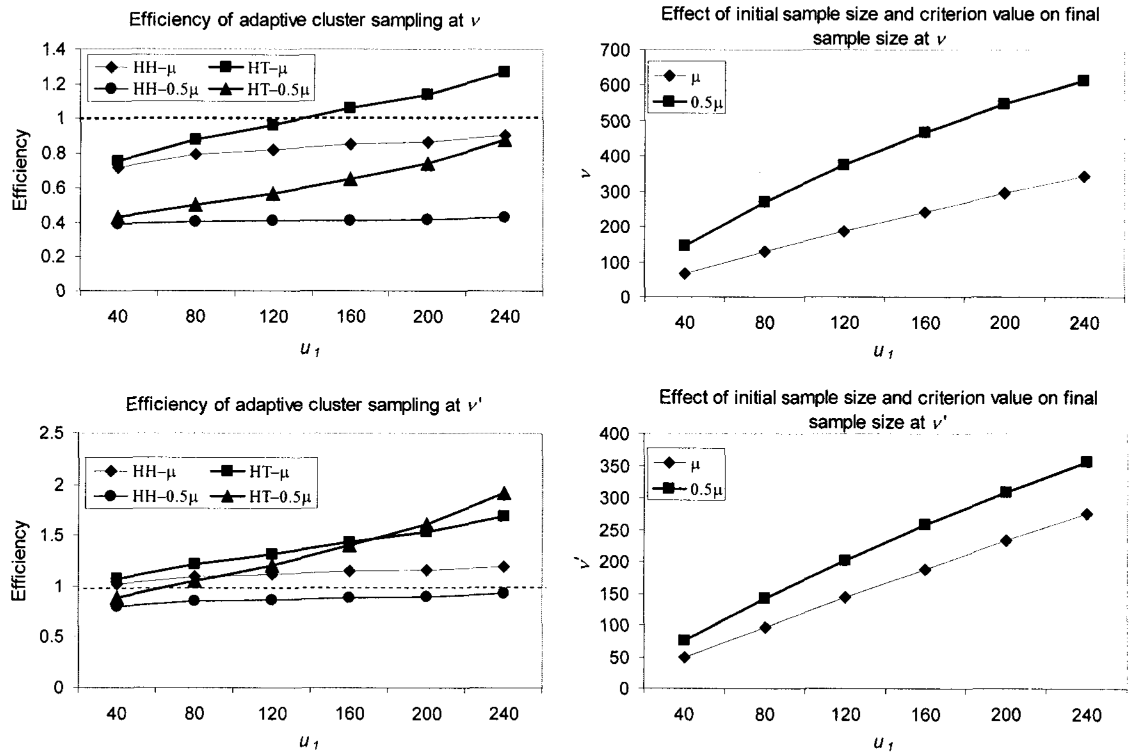


Figure 3.3. Efficiency of adaptive cluster sampling versus simple random sampling for a "highly aggregated" population (U=1600) from Su and Quinn (2003). *HH-μ* is the Hansen-Hurwitz estimator at the μ criterion level. *HT-0.5μ* is the Horvitz-Thompson estimator at the *0.5μ* criterion level. Sample size u_1 is the initial random sample size, v is final sample size including edge units, and v' is final sample size excluding edge units. Efficiency is relative to simple random sampling where the dashed line means equally efficient (1 or equal variances).

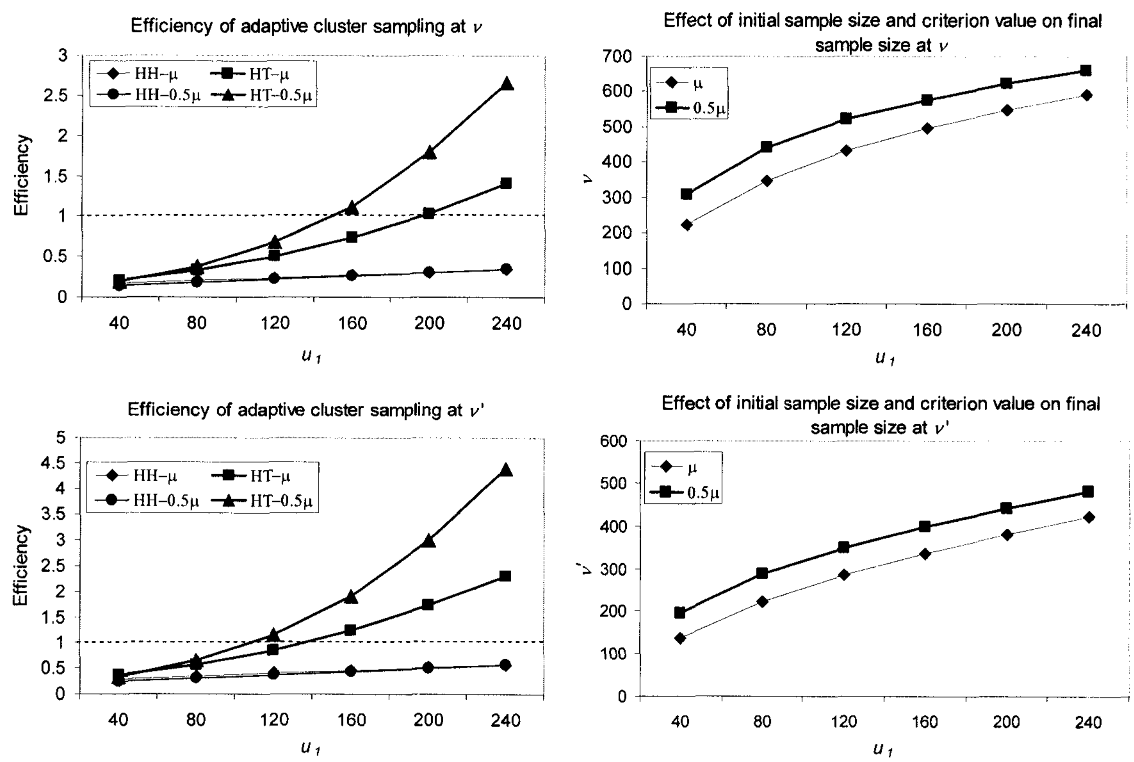


Figure 3.4. Adaptive sampling simulation results for an *S. alutus*-like simulated population (*U*=1600). *HH-μ* is the Hansen-Hurwitz estimator at the μ criterion level. *HT-0.5μ* is the Horvitz-Thompson estimator at the 0.5μ criterion level. Sample size u_1 is the initial random sample size, v is final sample size including edge units, and v' is final sample size excluding edge units. Efficiency is relative to simple random sampling where the dashed line means equally efficient (1 or equal variances).

Table 3.1. Summary statistics of two simulated populations. Populations were created in 40x40 grids (U=1600) with a Poisson cluster process (Diggle 2001). μ is the population mean, CV is the coefficient of variation and Pzr is the proportion of zeros cells in the population.

Population	μ	CV	Pzr
Hi aggregated[1]	190.6	3.8	0.84
S. alutus based[2]	191.1	4.3	0.16

[1] Based on "highly aggregated" population in Su and Quinn (2003).
[2] Used CPUE data from 2001 biennial survey between 150-300 m, scaled to equal the mean of population one.

3.5 References

Brown, J.A. 2003. Designing an efficient adaptive cluster sample. Environmental and Ecological Statistics 10:95-105.

Conners, M.E. and S.J. Schwager. 2002. The use of adaptive cluster sampling for hydroacoustic surveys. ICES Journal of Marine Science 59:1314-1325.

Diggle, P.J. 2001. Statistical Analysis of Spatial Point Patterns. 2nd edition, Oxford University Press, London.

Hanselman, D.H., T.J. Quinn II, C. Lunsford, J. Heifetz and D.M. Clausen. 2001. Spatial implications of adaptive cluster sampling on Gulf of Alaska rockfish. In Proceedings of the 17th Lowell-Wakefield Symposium: Spatial Processes and Management of Marine Populations, pp. 303-325. Univ. Alaska Sea Grant Program, Fairbanks, AK.

Hanselman, D.H., T.J. Quinn II, C. Lunsford, J. Heifetz and D.M. Clausen. 2003. Applications in adaptive cluster sampling of Gulf of Alaska rockfish. Fishery Bulletin 101:501-512.

Su, Z. and T. J. Quinn. 2003. Estimator bias and efficiency for adaptive cluster sampling with order statistics and a stopping rule. Environmental and Ecological Statistics 10:17-41.

Thompson, S.K. and G.A.F. Seber. 1996. Adaptive Sampling. Wiley, New York.

4 Utility of hydroacoustics to improve survey sampling precision for rockfish[9]

4.1 Introduction

Hydroacoustics is a technique used in fisheries and fisheries research that uses sound pulsed through the water to detect organisms in the water column as well as bottom depth and structure. This is done by using an echo-sounder which measures the sound that is reflected back to a transducer. The transducer is the physical unit attached to the vessel or a towed sled that sends and receives the sound pulses. The properties of sound transmitted through water are affected by water temperature, salinity, depth, and many other properties. The returned sound is then usually evaluated with software that adjusts for these different properties to obtain measures of organisms in the water column. These measures can vary from general statistics such as the average amount of backscatter per unit volume to the tracking of individual schools and fish.

The first successful experiment to detect fish acoustically was over seventy years ago (Kimura 1929). The technology has improved significantly, but the concepts and techniques remain relatively similar (MacLennan and Simmonds 1992). One notable improvement was the introduction of echo-integration (Dragesund and Olsen 1965), which is still widely used today to estimate absolute abundance in a number of fisheries. Echo-integration is basically summing the intensity of the backscatter for a given volume of water divided by the number of pulses sent out by the echosounder. Newer technology such as split-beam and dual-beam transducers (two pulses at the same frequency with a different size swath or two beams of different frequencies, respectively) has allowed

---

[9] Part 2 of 2 of: Hanselman, D.H. and Quinn, T.J. II. In Press Sampling Rockfish Populations: Adaptive Sampling and Hydroacoustics. *In* Sampling Rare or Elusive Populations: Challenges and Choices edited by William Thompson. Island Press. Split up by request of committee.

easier identification of specific fish in the field without prior laboratory experiments, and made calibration between surveys a simple task (MacLennan and Simmonds 1992).

Hydroacoustics has been used extensively by trawlers to locate rockfish (Major and Shippen 1970), but they have seldom been used for rockfish stock assessment. Some topics that hydroacoustics have been applied to rockfish research include species differentiation (Richards et al. 1991), exploratory surveys for biomass estimation (Kieser et al. 1993), and quantifying above-bottom schools (Starr et al. 1996). Recently, Stanley et al. (1999a) showed using hydroacoustics that yellowtail rockfish have nocturnal dispersion, but diel behavior did not significantly impact biomass estimates. New rockfish research on the Canadian Pacific coast has concentrated on estimating biomass in areas of high widow rockfish abundance (Stanley et al. 1999b, 2000, 2002). In these studies, they were able to use hydroacoustics to estimate abundance and delineate the size of a high density cluster for comparison with estimates from fishers. A key difference between Pacific ocean perch and widow rockfish is that Pacific ocean perch can be found in more trawlable topography, and are usually closer to bottom than widow rockfish (Brodeur 2001, Stanley et al. 2000, Zimmerman 2003). For Pacific ocean perch in the Gulf of Alaska, Krieger et al. (2001) showed a relatively strong relationship between catch rates and raw acoustic signal in a small study area. However, none of these studies attempted to apply hydroacoustics to improve large-scale sampling designs for rockfish. In this study I use hydroacoustic data collected from two years of NMFS bottom-trawl surveys to investigate how to improve survey precision for Pacific ocean perch.

4.2 Materials and Methods

Hydroacoustic data was recorded on Simrad® ES-60 (Simrad, http://www.simrad.com) echo-sounders on three vessels equipped with 38 KHz single-beam transducers. The vessels used in the Gulf of Alaska in 2001 were the F/V Vesteraalen (38 m stern trawler) and the F/V Morningstar (45 m stern trawler). The vessels used in the Aleutian Islands and Bering Sea in 2002 were the F/V Morningstar and the F/V Sea Storm (38 m stern trawler). Data during summer 2001 were recorded only during trawl hauls, whereas data recorded in summer 2002 were generally recorded continuously from before the first haul of the day until after the last one was completed. Nets, which were standard survey gear for the Gulf of Alaska, Aleutian Islands and Bering Sea, were towed an average of 15 minutes per haul (Martin and Clausen 1995). I used catch data from the National Marine Fisheries Service's RACEBASE (Oracle database containing all historical survey data) which included the catch weight, composition, numbers, location, time, distance fished, depth, and net dimensions.

I analyzed the raw hydroacoustic data with SonarData's Echoview® software (SonarData, http://www.sonardata.com), which I calibrated to the settings of each vessel. I analyzed all hauls from both years that contained any *S. alutus* and all hauls in the primary depth stratum for *S. alutus* (150-300m). The hydroacoustics were matched with each bottom trawl with a combination of GPS readings and time-stamps. I approximated the distance of the net behind the boat by using basic geometric methods (Pythagorean Theorem, ignoring wire curvature and horizontal angle). The track was then echo-integrated from 0.5m off bottom to the average net height measured by the net sounder (Scanmar). By using this small offset from the bottom, I may have included some of the "acoustic dead zone" (Stanley et al. 2000), the area close to bottom that fish become inseparable from bottom signal. However, I considered this offset to be a sensible choice considering that the mean net height is all within this possible dead zone. This choice may be reasonable considering the NMFS groundfish survey trawls on bottom topography that is relatively smooth and the weather in the summer minimizes pitch and roll of the vessel. This combination of factors should minimize this dead zone. The echo-integration resulted in a number of different measures of acoustic backscatter, but the mean volume backscattering (*Sv*) proved to have the best relationship with *S. alutus* CPUE (kg/km^3). I then used this variable to form a predictive model to convert *Sv*s to predicted catches for the sampling designs described below.

4.2.1 Analytical Methods and Sampling Designs

I computed 352 CPUE values (kg/km^3) from the three vessels and two years. The CPUE data were highly skewed and required transformation. I chose the flexible Box-Cox transformation which is a power transformation that uses maximum likelihood theory to estimate the optimum transformation. The distribution of these transformed values is shown in Figure 4.1. I compared these CPUE values to their respective *Sv*s from echo-integration. The data used is summarized in Table 4.1.

I first regressed these transformed CPUE values against *Sv* (Sokal and Rohlf 1995). I then examined sets of nonlinear models that added variables including depth, longitude, and catch composition. Depth can be a strong predictor for groundfish because many species occupy a specific depth interval. I selected longitude as a possible predictor because of density changes from west to east in the Gulf of Alaska. Catch composition was chosen as a possible variable because it could be used to differentiate areas of hydroacoustic signal with high densities of other species from areas with high densities of *S. alutus*. None of these variables would have an effect on the hydroacoustic signal, but would help predict when the signal may be representing *S. alutus.* I then selected the predictive model that yielded the best fit to the data in terms of R^2. I fitted each vessel's data to the model separately first, but confidence intervals for parameters overlapped across year and vessel; therefore, the final model was fitted with all the data pooled.

Because continuous hydroacoustic data were collected only in the 2002 surveys, these data were used for testing new sampling designs. For the 2002 data I examined a total of 182 hauls and their respective hydroacoustic tracks to be used in several sampling designs.

For the first design, I used the hydroacoustic model predictions to simulate a linear adaptive cluster sample (Thompson 1990) around the tows that exceeded a predetermined criterion value. For the initial sample (u), I used data from the 84 trawl hauls that contained *S. alutus* in the Aleutian Islands/Bering Sea data collected on the F/V Morning Star in 2002. I used the mean CPUE of the F/V Sea Storm from the same year as the criterion value to invoke ACS. I have performed simulations which showed that in an aggregated population, mean CPUE tends to be a reasonable criterion value that provides an improvement in precision with a relatively small final sample size. Additionally, by taking this value from the other vessel, it is independent of the sample (Hanselman et al. 2003).

I used the resulting adaptive networks to estimate mean abundance for the area sampled by the vessel. I used both the Hansen-Hurwitz-type (HH) and Horvitz-Thompson-type (HT) estimators (Thompson and Seber 1996) and compared them to standard simple random sampling (SRS) estimation. Smaller standard errors or coefficients of variation for the adaptive sampling estimators would indicate that adding the hydroacoustic model predictions resulted in an improvement in precision. I used nonparametric bootstrapping to estimate model variance and include this in the final variance. Equations are shown in Box 4.1.

The second design used to incorporate hydroacoustics into abundance estimation was double sampling (Thompson 2002). In double sampling, more precise estimates can be gained by using the relationship between an auxiliary variable that is easy to collect and the variable of interest that is more expensive or time-consuming. In this design, I used the observed tows as the variable of interest and used a larger sample of hydroacoustic model predictions of CPUE as the auxiliary variable. If the two variables have a high correlation and an intercept at the origin, then a ratio estimator can lead to a dramatic improvement in precision with negligible bias (Thompson 2002).

For this study, I sampled the Aleutian Island data from the F/V Sea Storm in 2002. A subsample of 40 tows and their respective hydroacoustic tracks was taken with another random sample of 80 hydroacoustic tows in the vicinity of these forty. This was an allocation of two auxiliary samples per trawl tow. The position of hydroacoustic tows were generated with uniform random numbers between -10 and 10 km away from each original station which was then set to be the center of a hydroacoustic tow of equivalent length. The model predictions from these tows were then combined in a ratio estimate of mean abundance and compared with the SRS estimates. The form of the estimators can be found in Box 4.1 with a simple example showing the gain in precision from using a ratio estimator with auxiliary data.

Hydroacoustic data may be useful for optimizing stratification of a survey. I show an example of using raw acoustic backscatter (*Sv*) as a basis for pooling the strata from the 2001 biennial groundfish survey conducted by NMFS. I use data from the Western and Central Gulf of Alaska (the Eastern Gulf was not surveyed in 2001). The original design used 59 strata to attempt to minimize variance for all species. Twenty-eight of these strata contained *S. alutus* in the 2001 survey. I used the hydroacoustic model predictions from 2001 for this area to pool these strata into four larger strata assigned by the four quartiles of the hydroacoustic predictions. No data were used between tows. The combined strata were used to determine the overall mean and variance of the abundance estimates using standard stratified-random sampling estimators (Thompson 2002). Although this method could be biased because the sampling design was originally stratified under an optimal allocation design, it illustrates the possible utility of using hydroacoustics as a means for stratification.

4.3 Results

When the transformed CPUE values were directly regressed on *Sv* (Figure 4.2), the coefficient and intercept were significant, but the fit was poor ($R^2 = 0.12$). The best model chosen to predict CPUE with hydroacoustics related the transformed CPUE to the natural log of *Sv* and the localized catch composition. The model took the form:

$$CPUE^* = a\ln(-Sv)^b + (COMP)^c$$

where $CPUE^*$ is the predicted Box-Cox transformed CPUE value with $\lambda = -0.05$ and $L = 2500$ (Sokal and Rohlf 1995), *Sv* is the mean volume backscattering value and *COMP* is the percentage of *S. alutus* in the closest tow. I used the pooled data (u=352) from all three vessels for both years. I estimated parameters using nonlinear least squares and obtained estimates of $a = 0.0539$, $b = 12.12$ and $c = 0.075$ ($p<0.0001$) with an

$R^2 = 0.82$ (Figure 4.3). I used this model to generate predictions for the following sampling design results.

4.3.1 Adaptive cluster sampling (ACS)

The ACS design allows for any number of patterns for adding additional samples as long as they are symmetric. I conducted the survey in a nearly linear pattern since the vessel moved in a transect from one sampling location to the next (see Figure 4.4). I added normal measurement error to the *COMP* variable of adaptive model predictions with mean equal to the *COMP* value of the tow with a CV of 0.33. This additional error simulated within-network variability more appropriately than model error alone and allowed *COMP* to vary naturally as the survey moved away from the original tow.

Previous work on adaptive sampling data for *S. alutus* showed that the sill of variograms (a measure of pairwise correlation of CPUEs with distance) produced for high abundance strata was approximately 10 km, roughly equating to average cluster size (Hanselman et al. 2001). Hence, I added adaptive samples linearly with model predictions until either they dropped below the criterion value or when they were more than 10 km away from the original tow.

Six tows of 84 exceeded the criterion value and the hydroacoustic model was used to add an additional fifty-five samples around these networks. Networks were bordered both by units not exceeding the criterion and by the 10 km distance limit. For this data set, there were lower estimates of mean abundance for ACS than the SRS estimators and fairly large gains in precision for both adaptive estimators (Table 4.2). If the same number of random samples were taken as those included in the adaptive estimator, SRS yielded a CV closer to the adaptive estimators (36%). However, the hydroacoustic samples required no extra ship time.

4.3.2 Double sampling

A requirement for the ratio estimator to perform well is that the two variables being used have a correlation near one and the variables are linearly related with an intercept at the origin. The subsample of 40 tows (y_i) and their respective 40 hydroacoustic model predictions (x_i) were well correlated ($r = 0.9$) with an intercept near the origin. The estimate of the ratio between the two variables was 1.07. When I used the 80 additional hydroacoustic predictions (x_i) in the ratio estimator, it performed efficiently. However, if the model estimates are back-transformed instead, the relationship becomes nonlinear and the estimator performs poorly compared to SRS. An example of some of the stations and their auxiliary hydroacoustic predictions are shown in Figure 4.5. The results from ratio estimation on the transformed scale are quite promising with 95% confidence intervals roughly half as wide as SRS (Figure 4.6). However when the estimation is done on the transformed scale and transformed back, it is no longer the arithmetic mean, but the geometric mean. While this is a valid statistic, it makes it more difficult to use in the standard stock assessment procedures employed by NMFS for rockfish, but could be useful as a biomass index.

4.3.3 Stratification

The pooled stratification assigned the twenty eight strata into four new strata (Table 4.3), with stratum one being the bottom quartile of model predictions and stratum four being the top quartile of model predictions. These pooled strata are not contiguous and would be difficult to show on a map, given the size of the Gulf of Alaska and the relatively small size of some of the strata on the continental slope.

The pooling of strata resulted in 4 strata containing 8 of the old strata from the Western Gulf and 20 from the Central Gulf. The Western Gulf appeared to contain more of the moderate strata, whereas the Central Gulf contained the very high and low densities. The depth distribution showed that the strongest hydroacoustic signal corresponded with the inclusion of the 100-300 meter depth stratum from the survey for strata 2-4. This depth range is where the density of *S. alutus* is highest (Hanselman et al. 2001). Stratum 1 looks much like stratum 4 in terms of areas and depths pooled because of the aggregation of the species. These areas may have been geographically proximate, but *S. alutus* simply were not encountered. This pooled stratification resulted in a substantial improvement in the precision of the estimate with only a minor change in the point estimate (Figure 4.7). Although this method may be biased because it is pooling strata of different sizes, the point estimate changes less than 4% between stratifications. This small change in the point estimate illustrates that the hydroacoustic data aided in gaining precision without much bias.

4.4 Discussion

The addition of hydroacoustic data into sampling designs for *S. alutus* showed promising results. The data were not collected randomly, but the data used in the analysis were subsampled randomly from the whole data set, allowing the assumptions of random sampling theory to apply. Since the additional data did not require additional ship time, and only the time of one analyst, the use of hydroacoustics may be an efficient way to gain precision.

I needed a model to use raw backscatter to predict CPUE values from the untrawled sections of the survey. The model I developed had an excellent fit to the data but may have been hyper-stable (little variability in predicted catch) because it used a localized catch composition to fit the data well. This type of model is more appropriate for use in double sampling or stratification than it is for the adaptive sampling design, which requires "true" samples to be unbiased. Double sampling and stratification do not have this restriction and auxiliary samples can simply be an "eyeball" estimate or any correlated variable that permits many samples to be collected easily.

Using the hydroacoustic data as a way to adaptively sample around units without actually performing tows is appealing. The unbiased estimator, however, intended by the original design likely has a bias of unknown size using the model-based predictions. Adaptive cluster sampling is most efficient when the within-network variability is high. In this example, the measurement error added into the *COMP* variable resulted in relatively high within network variability. Therefore, adding hydroacoustic samples linearly around tows showed substantial gains in precision but the estimate was lower than random sampling. This might be because of bias, or it could be similar to the previous applications of ACS to rockfish (Hanselman et al. 2001, Hanselman et al. 2003), where adaptive sampling yielded smaller estimates of mean abundance than SRS. However, this bias cannot be determined since true abundance is unknown. The validity of a model that uses catch composition becomes questionable the further away from the original tow.

In double sampling, the hydroacoustic model is not required to be accurate, just highly correlated with the variable of interest. In this example, this was true and the gains were substantial with a reduction of about one-half in the width of the 95% confidence intervals. Although a ratio estimate is not design-unbiased (its expected value does not equal the mean), a small bias is usually worth the trade-off for a large gain in precision. Double sampling with the untransformed data resulted in a less precise estimate than simple random sampling (SE 20% greater). Transformation of the data was necessary for the data to work well with the ratio estimator. Therefore, the problem with the use of a geometric mean in stock assessment remains. Taylor (1986) suggested several estimators for obtaining an arithmetic mean from the geometric mean of a Box-Cox transformation, but these estimators did not work well with these data because λ is near zero in the transformation (i.e., highly skewed). The double sampling estimates would work fine as an index. However, for an index to be useful, a time series of measurements is needed, where any bias is relatively constant. I have collected data for two surveys (2001, 2003) in the Gulf of Alaska and should continue to collect this data as it may become more useful as I accumulate a longer time series.

Using the hydroacoustic model for stratification showed some potential utility. My heuristic approach of pooling strata of different sizes is statistically biased, but changed the abundance estimate very little while reducing the confidence intervals substantially. Lunsford (2001) showed that most of the gains in precision in the current stratified design were from allocation, not from the strata chosen indicating that the current design was not proficient at separating low and high density POP areas. This indicates that hydroacoustics may be useful for stratification of future designs for rockfish, or real-time stratification. One application of this type is to stratify on-the-fly by allocating more tows to patches of higher density hydroacoustic signal and fewer tows to areas of low-density signal as the survey progresses. This study was illustrated with mackerel icefish (*Champsocephalus gunnari*) (Everson et al. 1996) and experiments are in progress to test the methods on rockfish (Paul Spencer pers. comm.[10]).

4.4.1 Conclusions and future work

Future work for sampling *S. alutus* should concentrate on obtaining more samples within their habitat range, perhaps with a vessel equipped with a more rugged net to assess areas that previously only hydroacoustics have sampled. This would allow a more random and representative sample of *S. alutus* habitat and their distribution than the current gear allows. Research on utilizing the low-cost and readily available hydroacoustic information that is currently being collected on survey vessels should be continued. Collecting hydroacoustic data from commercial vessels throughout the season is also a possibility. Further comparison of what the net is catching versus what the echo-sounder is recording should be done by submersibles or towed sleds to validate the use of hydroacoustic signal so close to the bottom. The most promising use of the data thus far is in a double sampling design. Future work with these data should concentrate on how to use this as an index in the current stock assessment model and on using hydroacoustic data for stratification of future surveys.

4.5 Acknowledgments

I would like to thank Jeff Fujioka at the Auke Bay Laboratory who originally started looking at hydroacoustic data for rockfish. I also would like to thank the group at the Alaska Fisheries Science Center for their help in collecting the data, specifically Paul Spencer, Bill Flerx, Jerry Hoff and the other field party chiefs in the 2001 and 2002

[10] Paul Spencer. National Marine Fisheries Service. Alaska Fisheries Science Center. 7600 Sand Point NE, Seattle WA.

survey seasons. Finally, I thank the captains of the F/Vs Morningstar, Vesteraalen and Sea Storm for allowing me to tinker with their echo-sounders.

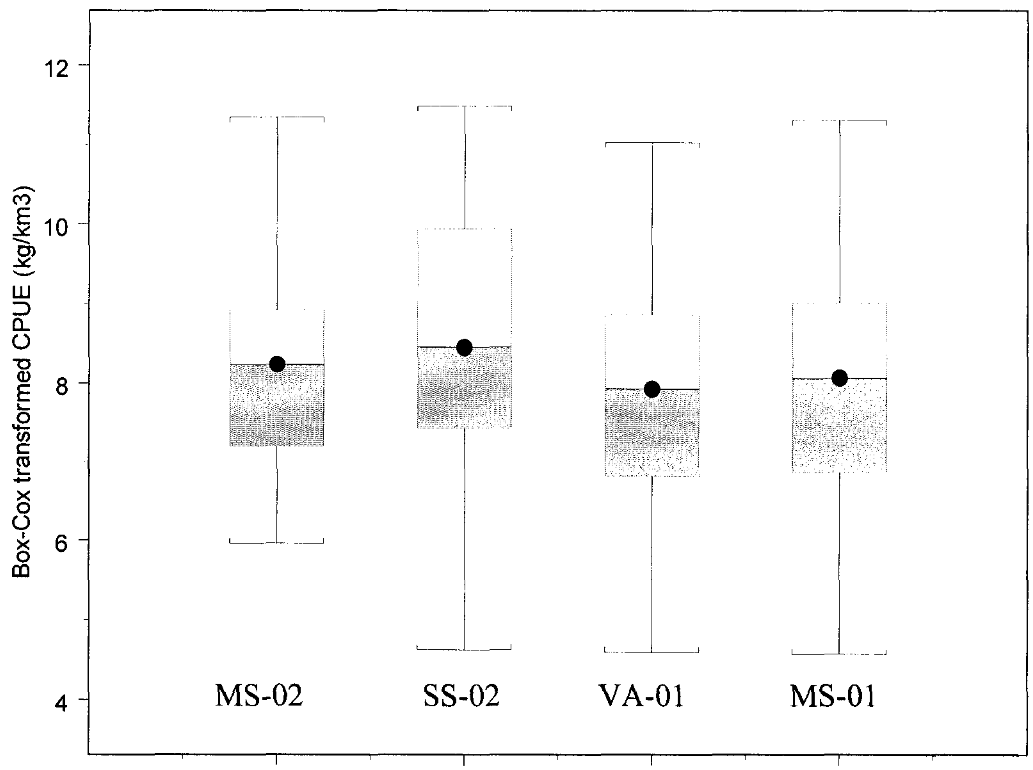


Figure 4.1. Boxplots of Box-Cox transformed CPUE (kg/km3) values for three vessels and two years of the NMFS groundfish surveys. MS02=F/V MorningStar 2002, SS02=F/V Sea Storm 2002, VA01=F/V Vesteraalen 2001, MS01=F/V MorningStar 2001. The box contains the middle half of the data, the line and point in the box represent the median, and the whiskers represent 1.5x(interquartile range).

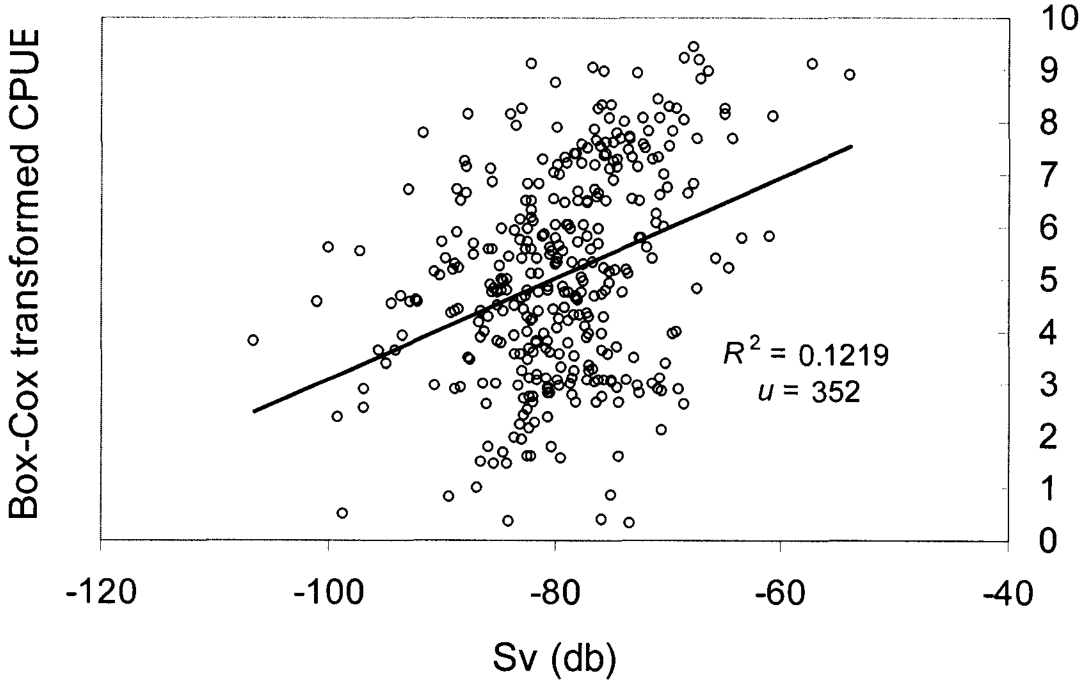


Figure 4.2. Fit of Box-Cox transformed Catch per unit effort (CPUE, kg/km^3) of Pacific ocean perch versus mean volume backscatter (Sv) for the F/V Sea Storm 2002.

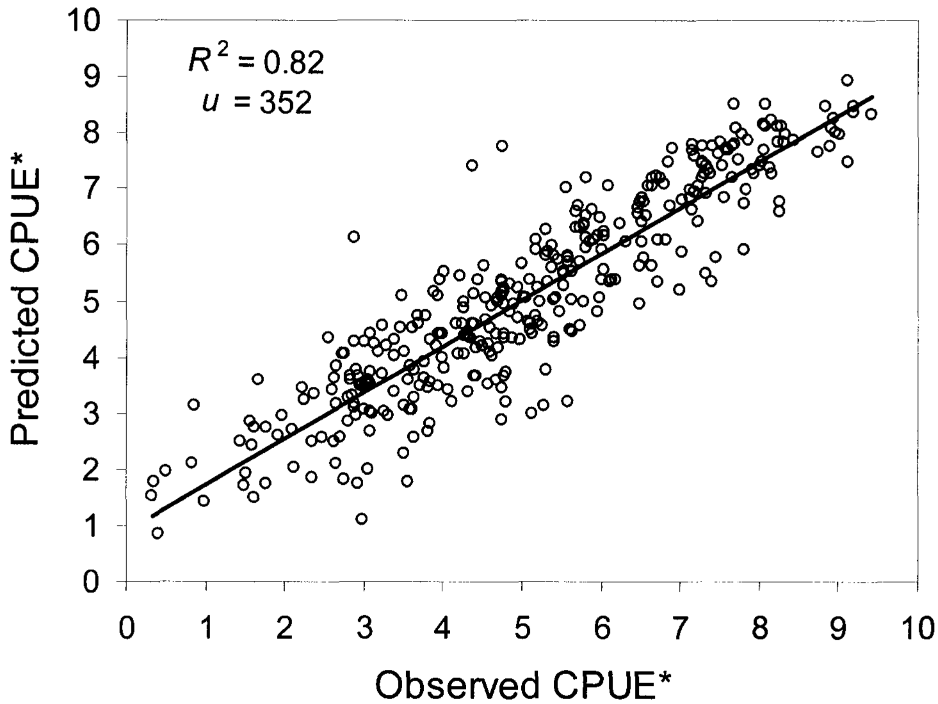


Figure 4.3. Plot of hydroacoustic model predicted CPUE* versus the observed transformed CPUE* (Box-Cox power transformed) for pooled data for all vessels, all years.

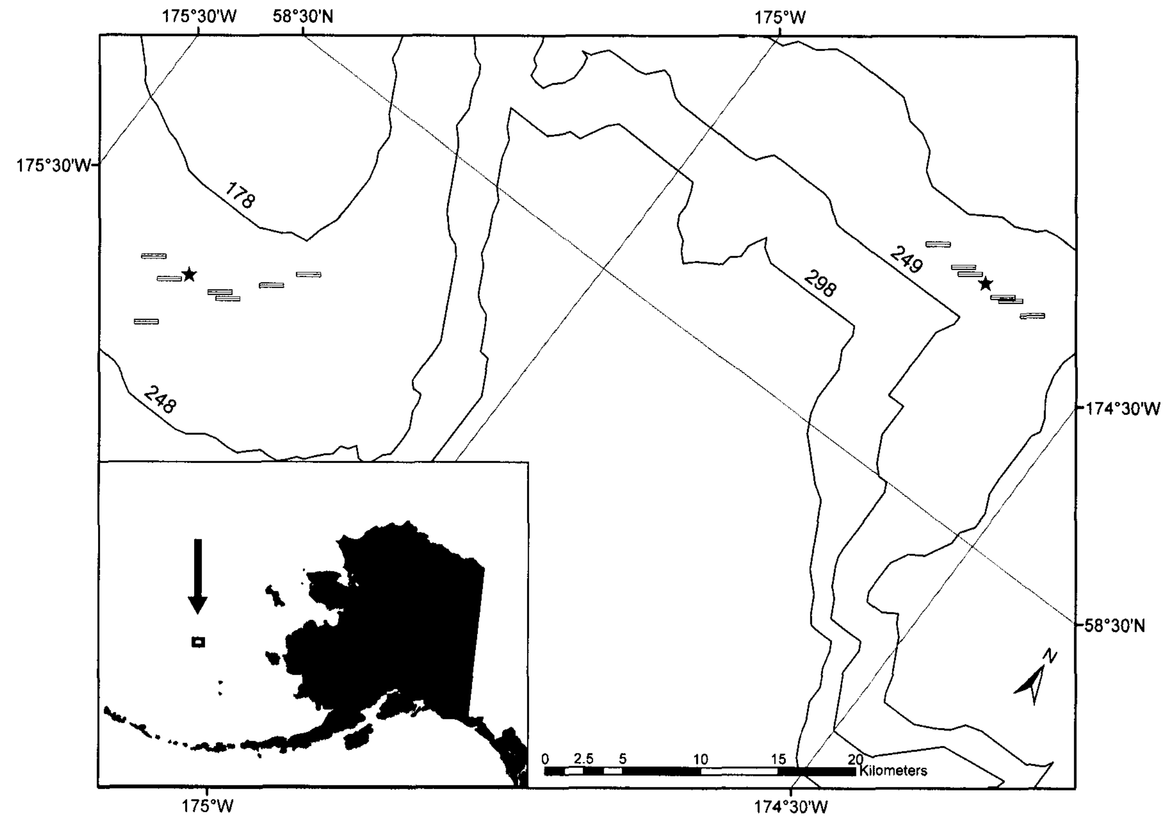


Figure 4.4. Location of two simulated adaptive cluster samples for Pacific ocean perch. Data were from F/V Morning Star 2002 in the Bering Sea. Stars are survey tows, the thick lines are the predicted tow tracks from hydroacoustic observations. Thin contour lines are bathymetry (in meters).

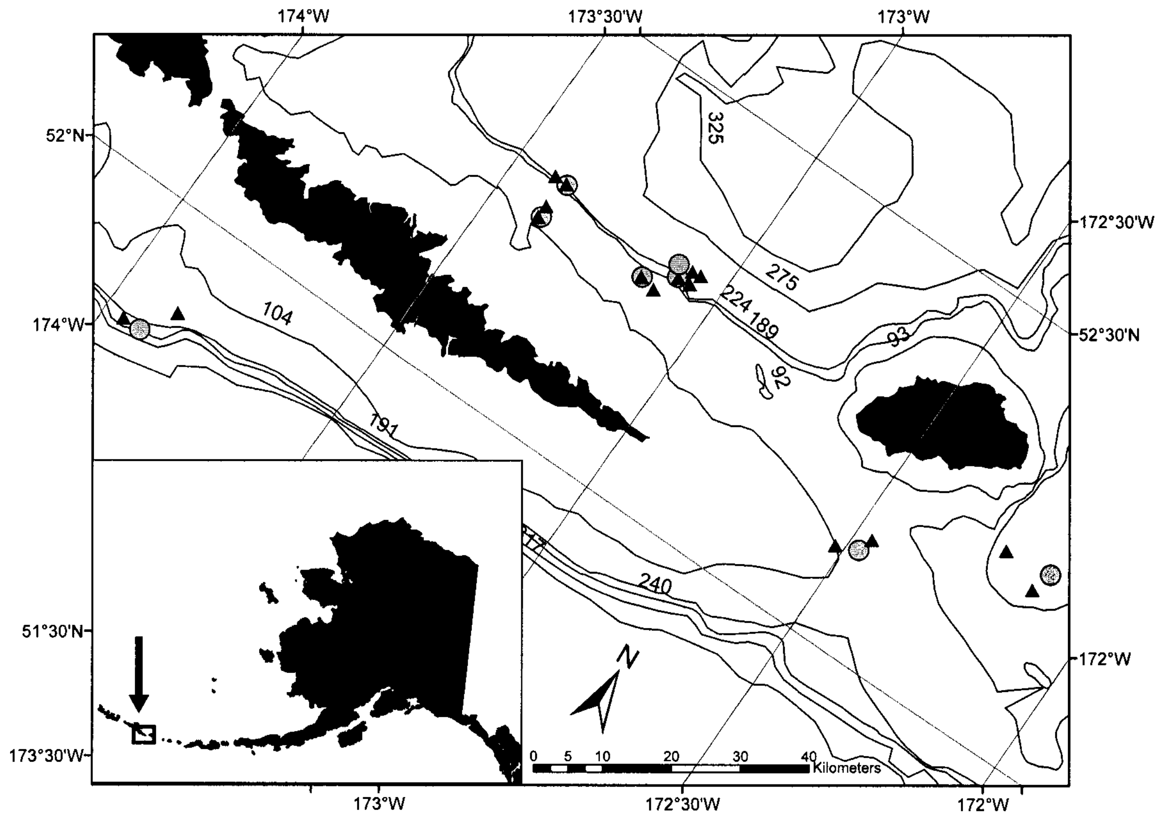


Figure 4.5. Map of some double sampling locations from the F/V Sea Storm 2002 in the Aleutian Islands. Grey circles represent the center of the tow sample. Black triangles represent the random acoustic model predictions for double sampling. Numbered contour lines are bathymetry (in meters).

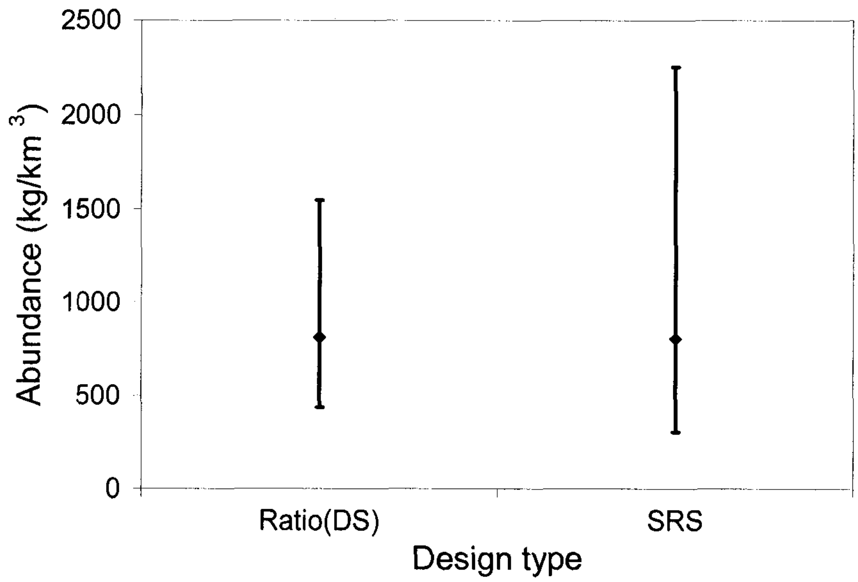


Figure 4.6. 95% confidence intervals for a ratio estimate (double sampling) and a simple random sampling estimate (SRS) for mean abundance of one ship's catch in the Aleutian Islands 2002.

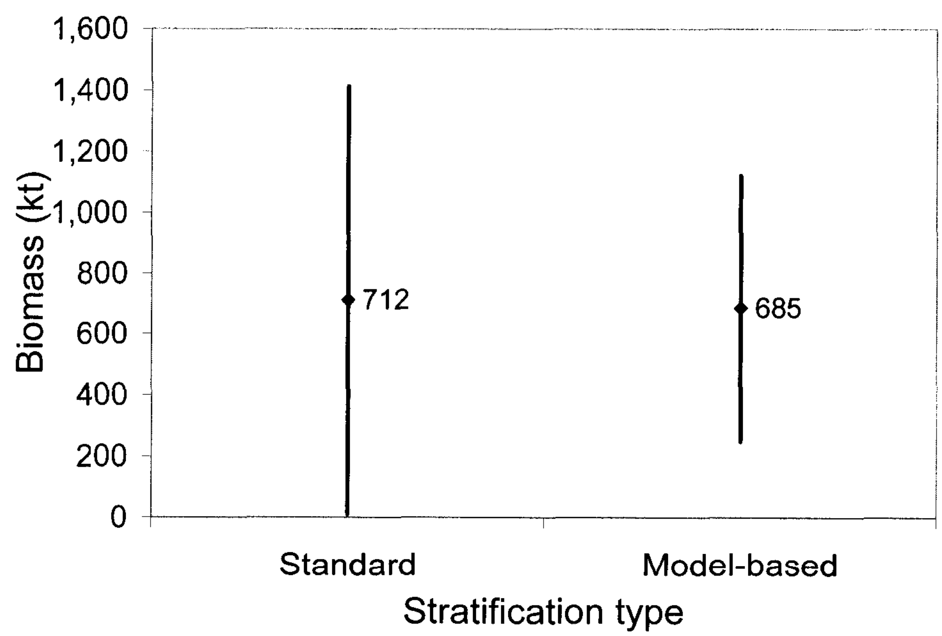


Figure 4.7. Comparison of using the original stratification of the 2001 NMFS biennial groundfish survey for S. alutus and using a hydroacoustic model-based approach. Error bars represent 95% confidence intervals. Data used was for the Western and Central Gulf of Alaska and hydroacoustic data was taken from 2001 F/V Morning Star and the 2001 F/V Vesteraalen.

Table 4.1. Data collected from different vessels, years and areas for hydroacoustic analysis of S. alutus catches. Continuous means that hydroacoustics were recorded between tows, sample size refers to the number of tows with S. alutus catch analyzed.

Data Set	Vesteraalen 2001	Morning Star 2001	Morning Star 2002	Sea Storm 2002
Area	Gulf of Alaska	Gulf of Alaska	Aleutian Islands/Bering Sea	Aleutian Islands
Continuous?	No	Some	Yes	Yes
Sample size	83	87	85	97

Table 4.2. Results of a hydroacoustic adaptive sampling experiment. HT=Horvitz-Thompson adaptive estimator, HH=Hansen-Hurwitz adaptive estimator, SRS= Simple random sampling estimator, $\hat{\mu}$ =estimate of mean abundance (kg/km), SE=standard error, CV=coefficient of variation. SRS-ν is simple random sampling at the same sample size as adaptive sampling.

Estimator	SRS	SRS-ν	HH	HT
$\hat{\mu}$	10773	10773	5912	6086
SE	5042	3878	1683	1681
CV	47%	36%	28%	28%

Table 4.3 Results of pooling strata for 2001 NMFS groundfish survey (Eastern Gulf was not surveyed in 2001). Strata are ranked in quartile of hydroacoustic density with 4 being the highest. Western is number of strata from the Western Gulf and Central is number of strata from the Central gulf. 100-300 m depth is the percentage of strata that are in a depth between 100-300 m.

Stratum	# of old strata	Combined area (km^2)	Average Density (kg/km^2)	Western	Central	100-300 m depth
1	7	48,615	0.57	1	6	71%
2	7	58,246	2.59	3	4	29%
3	7	42,832	55.8	3	4	43%
4	7	43,752	98.0	1	6	86%
Total	28	193,445	35.4	8	20	39%

4.6 References

Brodeur, R.D. 2001. Habitat-specific distribution of Pacific ocean perch (*Sebastes alutus*) in Pribilof Canyon, Bering Sea. Continental Shelf Research 21:207-224

Dragesund, O. and S. Olsen. 1965. On the possibility of estimating year-class strength by measuring echo-abundance of 0-group fish. FiskDir. Skr. Ser. Havunders. 13:47-75. Cited in MacLennan and Simmonds (1992).

Everson, I., M. Bravington and C. Goss. 1996. A combined acoustic and trawl survey for efficiently estimating fish abundance. Fisheries Research 26:75-91.

Hanselman, D.H., T.J. Quinn II, C. Lunsford, J. Heifetz and D.M. Clausen. 2001. Spatial inferences of adaptive cluster sampling on Gulf of Alaska rockfish. Proceedings of the 17th Lowell-Wakefield Symposium: Spatial Processes and Management of Marine Populations, 303-325.

Hanselman, D.H., T.J. Quinn II, C. Lunsford, J. Heifetz and D.M. Clausen. 2003. Applications in adaptive cluster sampling of Gulf of Alaska rockfish. Fishery Bulletin 101:501-512.

Kieser, R.B., K. Cooke, R.D. Stanley, and G.E. Gillespie. 1993. Experimental hydroacoustic estimation of rockfish (*Sebastes* spp.) biomass off Vancouver Island, November 13-25 1991. Canadian Manuscript Report of Fisheries and Aquatic Sciences 2185.

Kimura, K. 1929. On the detection of fish groups by an acoustic method. J. imp. Fish Inst., Tokyo 24:41-5.

Krieger, K., J. Heifetz, and D. Ito. 2001. Rockfish assessed acoustically and compared to bottom-trawl catch rates. Alaska Fishery Research Bulletin 8-1:71-77.

Lunsford, C., L. Haldorson, J.T. Fujioka, and T.J. Quinn II. 2001. Distribution patterns and survey design considerations of Pacific ocean perch (*Sebastes alutus*) in the Gulf of Alaska. Proceedings of the 17th Lowell-Wakefield Symposium: Spatial Processes and Management of Marine Populations, 303-325.

MacLennan, D.N. and J.E. Simmonds. 1992. Fisheries Acoustics. Chapman and Hall, New York.

Major, R.L. and H.H. Shippen. 1970. Synopsis of biological data on Pacific ocean perch, Sebastodes alutus. FAO Fisheries Synopsis 79, and U.S. Dept. of Commerce, National Marine Fisheries Service NMFS/S 79, Circular 347.

Martin, M.H. and D.M. Clausen. 1995. Data Report: 1993 Gulf of Alaska bottom trawl survey. NOAA Technical Memorandum NMFS-AFSC-59.

Quinn, T.J. II. and R.B. Deriso. 1999. Quantitative Fish Dynamics. Oxford University Press, New York.

Richards, L.J., R. Keiser, T.J. Mulligan and J.R. Candy. 1991. Classification of fish assemblages based on echo integration surveys. Canadian Journal of Fisheries and Aquatic Sciences 48:1264-1272.

Sokal, R.R. and F.J. Rohlf. 1995. Biometry: the principles and practice of statistics in biological research. 3rd edition. Freeman, New York.

Stanley, R.D., A.M. Cornthwaite, R. Kieser, K. Cooke, G. Workman and B. Mose. 1999a. An acoustic biomass survey of the Triangle Island widow rockfish (*Sebastes entomelas*) aggregation by Fisheries and Oceans, Canada and Canadian Groundfish Research and Conservation Society, January 16-February 7, 1998. Canadian Technical Report of Fisheries and Aquatic Sciences 2262.

Stanley, R. D., R. Kieser, B.M. Leaman and K. Cooke. 1999b. Diel vertical migration by yellowtail rockfish (*Sebastes flavidus*) and its impact on acoustic biomass estimation. Fishery Bulletin 97: 320–331.

Stanley, R. D., R. Kieser, K. Cooke, A. M. Surry, and B. Mose. 2000. Estimation of a widow rockfish (*Sebastes entomelas*) shoal off British Columbia, Canada as a joint exercise between stock assessment staff and the fishing industry. ICES Journal of Marine Science 57:1035–1049.

Stanley, R. D., R. Kieser and M. Hajirakar. 2002. Three-dimensional visualization of a widow rockfish (*Sebastes entomelas*) shoal over interpolated bathymetry. ICES Journal of Marine Science 59:151–155.

Starr, R.M., D.S. Fox, M.A. Hixon, B.N.Tissot, G.E. Johnson and W.H. Barss. 1996. Comparison of submersible-survey and hydroacoustic-survey estimates of fish density on a rocky bank. Fishery Bulletin 94:113-123.

Taylor, J.M.G. 1986. The retransformed mean after a fitted power transformation. Journal of the American Statistical Association 81:114-118.

Thompson, S.K. 1990. Adaptive cluster sampling. Journal of the American Statistical Association 412:1050–1059.

Thompson, S.K. 2002. Sampling. 2nd edition, Wiley, New York.

Thompson, S.K. and Seber, G.A.F. 1996. Adaptive Sampling. Wiley, New York, 265 p.

Zimmerman, M. 2003. Calculation of untrawlable areas within the boundaries of a bottom trawl survey. Canadian Journal of Fisheries and Aquatic Sciences. 60:657–669.

Box 4.1. Equations for estimation of abundance for adaptive sampling and double sampling.

Estimation of Hansen-Hurwitz estimator in ACS using hydroacoustic adaptive units

$$\hat{\mu}_{\mathrm{HH}} = \frac{1}{u}\sum_{1}^{u}\frac{\hat{y}_i^*}{x_i} = \frac{1}{u}\sum_{1}^{u}\hat{w}_i^*;\ \text{and } s_w^2 = \sum_{1}^{u}\frac{\left(\hat{w}_i^* - \hat{\mu}_{\mathrm{HH}}\right)^2}{u-1}$$ from Thompson 2002 p.294

where $\hat{y}_i^*$ is the network total from an actual tow y_{i1} that exceeds a criteron value and its corresponding hydroacoustic samples $\hat{y}_{i2}, \hat{y}_{i3}, \ldots \hat{y}_{iu,y>c}$. $\hat{w}_i^*$ is the network mean

$\hat{y}_{ij} = a(Sv) + b(Comp)^c$ best-fit model from hydroacoustic data set, u=352,$R^2 = 0.82$, where SV = acoustic backscatter,$Comp$ = Catch composition for species

$\hat{\mathrm{var}}(\hat{y}_{ij}) = Sv^2\ \mathrm{var}(a)_{boot} + \mathrm{var}[b(Comp)^c]_{boot}$ from non-parametric residual resampling (Quinn and Deriso 1999)

$\hat{\mathrm{var}}(\hat{w}_i^*) = \sum_{2}^{x|y>c}\mathrm{var}(\hat{y}_{ij}) + \sum_{j=1}^{x}\mathrm{cov}(\hat{y}_{ij}, \hat{y}_{ij+1})$, where $\sum_{j=1}^{x}\mathrm{cov}(\hat{y}_{ij}, \hat{y}_{ij+1})$ is assumed equal to 0 (Sokal and Rohlf 1995, p.567)

$$\hat{\mathrm{var}}(\hat{\mu}_{\mathrm{HH}}) = \left(1 - \frac{u}{U}\right)\frac{s_w^2}{u} + \frac{1}{Uu}\sum_{i=1}^{u}\frac{\mathrm{var}(\hat{w}_i^*)}{x_i}$$ (Thompson 2002 two-stage estimator)

Estimation of ratio estimator by using hydroacoustic model predictions as an auxilary variable

$$\hat{r} = \frac{\sum_{1}^{u} y_i}{\sum_{1}^{u}\hat{x}_i}$$ from Thompson 2002 p.158 where $\hat{r}$ is the ratio of the sum of observed catches y_i to the sum of corresponding model predicted catches from hydroacoustics $\hat{x}_i$.

$\hat{\mu}_x = \frac{1}{u'}\sum_{1}^{u'}\hat{x}_i$ where u' is the number of auxiliary predictions $\hat{x}_i$

$\hat{\mu}_r = \hat{r}\hat{\mu}_x$ which is the estimate of the population mean

$\hat{\mathrm{var}}(\hat{\mu}_r) = \left(1 - \frac{u'}{U}\right)\frac{s_y^2}{u'} + \left(\frac{u'-u}{u'u(u-1)}\right)\sum_{i=1}^{u}(y_i - \hat{r}\hat{x}_i)^2$ it does not matter that $\hat{x}_i$ is estimated because it is auxiliary and could be compared to an 'eyeball' estimate

$\hat{x}_i = a(Sv) + b(Comp)^c$ best-fit model from hydroacoustic data set, u=352,$R^2 = 0.82$, where Sv = acoustic backscatter,$Comp$ = Catch composition for species

---

Box 4.1 (continued).

Example of using a ratio estimator with auxiliary data

$U = 100, u = 3, u' = 11$

$$x_i = \left[\overset{u'}{\overbrace{\overset{u}{(2,6,8)} : 4,7,2,5,7,8,10,12}} \right] \text{ auxiliary variable}$$

$y_i = (3,7,9)$ variable of interest

$s_y^2 = 9.33$ sample variance of y_i

$$\hat{r} = \frac{\sum_1^u y_i}{\sum_1^n x_i} = \frac{(3+7+9)=19}{(2+6+8)=16} = 1.19 \text{ ratio of subsample of } y_i\text{s and } x_i\text{s}$$

$$\hat{\mu}_x = \sum_1^{u'} \frac{x_i}{u'} = (2+6+8+4+7+2+5+7+8+10+12)/11 = 6.45 \text{ mean of auxiliary variable}$$

$$\hat{\mu}_r = \hat{r} \times \overline{x} = 6.45 * 1.19 = 7.7 \text{ ratio estimate of mean}$$

$$\hat{\text{var}}(\hat{\mu}_r) = \frac{(U-u')}{U}\frac{s^2}{u'} + \left[\frac{u'-u}{u'u(n-1)}\right]\sum_1^u (y_i - \hat{r}x_i)^2 \text{ variance of ratio estimate}$$

$$\hat{\text{var}}(\hat{\mu}_r) = \frac{(100-11)}{100}\frac{9.33}{11} + \left[\frac{11-3}{11\times 3(3-1)}\right] \times 0.66 = 0.83$$

$$\hat{\text{var}}(\hat{\mu}_{srs}) = \left(\frac{U-u}{U}\right)\frac{s^2}{u} = \left(\frac{100-3}{100}\right)\frac{9.33}{3} = 3.01 \text{ variance of SRS estimate}$$

5 Five simulation experiments using an age-structured assessment model: Survey error, data weighting and priors[11]

5.1 Introduction

Separable age-structured stock assessment or Integrated Analysis (IA) has become one of the most common and advanced methods for assessing stock status on the West Coast of North America (Fournier and Archibald 1982, Methot 1989, Quinn and Deriso 1999, Butterworth et al. 2003). One software product for constructing these models is AD Model Builder[12]. This commercial product provides a framework for general model building using an automatic differentiation algorithm to optimize an objective function. The flexibility of the framework allows for the incorporation of measurement and process error, full Bayesian estimation, and the ability to efficiently estimate hundreds of parameters. The tremendous flexibility of this type of modeling allows the stock assessment practitioner to explore many aspects of the modeling procedure, including both the effects of the data inputs and the more subjective inputs such as data weightings and prior distributions.

In stock assessment models that integrate multiple data sources, the effects of any one data source can be altered by varying components of the model. It is important to evaluate the behaviour of these components on model results such as estimates of biomass, harvest rate, and recruitment. These types of stock assessments have not been extensively evaluated through simulation. Bence et al. (1993) examined the influence of age-specific selectivity patterns under the Stock Synthesis model (Methot 1989). The National Research Council (1998) used several simulated data sets to test a number of popular stock assessment methods including the IA method, which performed well. Sampson and Yin (1998) examined eight factors under the Stock Synthesis model,

---

[11] Hanselman, D.H. and Quinn, T.J. II. *In prep.* Canadian Journal of Fisheries and Aquatic Sciences.

[12] Otter Research Ltd. P.O. Box 2040, Sidney, B.C. V8L3S3, Canada.

including natural mortality, recruitment variation and survey coefficients of variation (CVs). Punt et al. (2002) examined IA and five other methods to evaluate data from four different species, finding that IA generally performed best. Most other recent simulation studies have contrasted the IA approach with other methods. In this study, I focus on evaluating factors that affect the output of the IA approach. Two important factors in evaluating a stock assessment model are the effects of errors in the survey biomass indices, and the effects of using prior information and weightings to express knowledge about parameters and data.

In this study, I use a stock assessment model for Pacific ocean perch (*S. alutus*) in the Gulf of Alaska (Hanselman et al. 2003a) to explore uncertainties in components of the model. The results of the 2003 stock assessment model serve as the "true" values and simulated data sets are constructed in five experiments to answer the following questions: (1) What is the effect of measurement error in survey biomass estimates on stock assessment results? (2) What is the effect of changing catchability over time? (3) What does adding an additional biomass index do to model precision? (4) How sensitive are model results to different weighting schemes on the data sources? (5) How sensitive are model results to prior distributions imposed on key parameters?

5.2 Methods

The generating model was the annual stock assessment conducted for the Pacific ocean perch (*Sebastes alutus*) stock in the Gulf of Alaska (Hanselman et al. 2003a). The species is a long-lived fish, with a substantial historical fishery (350,000 mt in 1965). The population is beginning to recover from the period of heavy exploitation in the 1960s, and now has a modest annual trawl fishery (catch of ~12,000 mt in 2003). I used this stock assessment model to generate the "true" population. The model spanned a time period of 43 years (1961-2003) and included multiple sources of data, including fishery and survey ages (age distributions by year), fishery lengths (distribution of fish lengths from the fishery by year), catch by year, and a survey biomass index (Table 5.1). The model was a separable age-structured model similar to others conducted on the West Coast of North America for species such as walleye pollock. Key parameters estimated included catchability (q), natural mortality (M), selectivity by age for the survey and fishery (s_a), fishing mortality by age and year (f_{ya}), and recruitment deviations. Unlike similar models, this model did not attempt to fit a stock-recruitment relationship; rather it estimated a mean recruitment, which was adjusted by estimated recruitment deviations for each year. Equations of the simulation model were presented in Box 1.1 of Chapter 1.

5.2.1 Experimental design

Simulations were implemented with a dynamic link library (DLL) compiled with AD Model Builder. I constructed simulated data sets and passed them to the DLL from S-Plus 6 Professional. Five experiments were conducted (Table 5.2). For experiments 1-3, survey biomass indices were simulated with varying levels of measurement error. The assessment model was then run using each simulated biomass index with the assigned survey precision of that index equal to the measurement error added. This procedure assumes that the analyst correctly estimated the variance when the population was sampled. The survey precision assigned to the biomass estimates acts as a weighting term which determines how much influence the survey biomass estimates have on the overall model. These are not to be confused with the simulated level of measurement error or the precision of the stock assessment results. Simulations are also compared to a "base case", in which there was no measurement error in the index and a low survey CV estimate of 1% to represent that the index was measured with little error. This also showed that the model can reproduce its own results.

Experiment 1 examined the effects of varying levels of measurement error in the primary survey biomass index. I generated 200 data sets of 4 levels of measurement error in the biomass index (CV=5, 10, 25, 50%). I used 200 data sets in this experiment to test if 100 simulations were sufficient for the subsequent experiments. This resulted in 801 simulations (1 for the base case simulation where measurement error = 0 and 200 for each of the other four levels). I expected that the estimating model would produce estimates without substantial bias, but would produce imprecise estimates as measurement error increased.

Experiment 2 examined the effects of a trend in catchability over time. A positive trend in catchability could be explained by an increasing efficiency of the gear as the survey team and captains become more experienced. An example of a negative trend in catchability could occur as more prime rockfish habitat was deemed "untrawlable" substrate. I evaluated four levels of trend in the catchability: decreasing 5 and 10% per survey and increasing 5 and 10% per survey. One hundred sets of survey biomass estimates were generated for each trend with measurement error added (CV=5, 10, 25, 50%) for a total of 1600 simulations. I expected that if catchability decreased over time it would result in a modest downward bias in biomass estimates with a modest decrease in precision with increasing measurement error. With an increasing trend in catchability, I expected a large upward bias in biomass and a large decrease in precision with increasing measurement error. The biomass indices with a trend and high measurement error should have lower precision than the indices with low measurement error, but less bias.

Experiment 3 examined the effects of adding a second biomass index (Index 2) to the model. The index was modeled as an acoustic or environmental index, with no additional information such as ages and sizes. I simulated adding five additional biomass estimates corresponding to the last five years of surveys such as the acoustic index described in Chapter 4. The second index was weighted equally to the first index at one. I examined three trends in Index 2: A "confirming" trend, where the annual means of the secondary index are the same as Index 1, an index with a downward trajectory of 10% from Index 1, and an index with an upward trajectory of 10% from Index 1. One hundred data sets of each trend were constructed with measurement error CVs of 5, 10 and 25%. The higher CV of 50% was not used here because, presumably, I would be using a new index because it was more precise than the previous one. These data sets were matched with measurement error CVs in Index 1 of 10, 25, and 50%. This resulted in 900 simulations (combinations of CV=10/5%, 25/10%, and 50/25% with three trends). My expectation was that when the two survey indices were showing the same trend, bias should be low and precision would be better than the corresponding measurement error levels in Experiment 1. Estimated standard deviations (from AD Model Builder Hessian matrix) of the biomass estimates should be reduced. When the trends were contradictory, the results would be less precision and higher bias. Estimated standard deviations of the biomass estimates would be reduced less than the confirming trend.

Experiment 4 examined the effects of using deliberate weightings on data sources. For the five data sources in the base model, four had a weight of one (fishery ages, fishery lengths, survey ages and the survey biomass index) and one had a weight of 50 (catch, which was believed to be measured with little error). For these five data sources I conducted 25 simulations by adjusting one data source weighting while holding the other four equal to the base model. For catch I used weights of 5, 10, 50, 100, and 500. For the other four data sources I used weights of 0.1, 0.5, 1, 2, and 10. I expected that the weighting schemes would have a significant effect on both bias and precision when the weights of data sources that contributed most to the objective function value (the number that was minimized to provide the best fit to all model components) were changed. The percentage of each component in the objective function total of the generating model was: Survey Ages=43%, Fishery Sizes=32%, Fishery Ages=21%, Survey biomass=4% and Catch=<1%.

Experiment 5 examined the effects of using prior distributions on key parameters. In the base model I estimated three parameters using lognormal prior distributions: catchability $q \sim lognormal$ (μ=1,CV=0.2), recruitment variability $\sigma_r \sim lognormal$ (μ=1.7,CV=0.2) and natural mortality $M \sim lognormal$ (μ=0.05,CV=0.01). For this experiment I varied the means and CVs of these prior distributions to test the sensitivity of the model to these assumptions. For catchability and recruitment variability, I tested five means of (0.5, 0.75, 1, 1.33, and 2) and (0.85, 1.275, 1.7, 2.27 and 3.4) respectively with CVs of 0.1, 0.2, and 0.4. Natural mortality was a very sensitive parameter and difficult to estimate, so it has a precise prior. I examined natural mortality means of 0.025, 0.0375, 0.05, 0.067, and 0.1 with CVs of 0.005, 0.01 and 0.02. This was done for a total of 45 (3x5x3) simulations.

5.2.2 Evaluation of effects

I evaluated the effects of different errors on the estimation ability of the model by examining the distribution of relative errors to the "true" population. I used the coefficient of variation of each set of simulations as a measure of relative estimation precision (REP). A high REP would be indicated by a low coefficient of variation. I also examined relative estimation accuracy (REA) or bias caused by each factor as median relative error, $100x(median(\hat{\theta}_y)-\theta)/\theta$) of each parameter from the parameters of the "true" population. Median unbiasedness, when REA was close to zero, would indicate that overestimates and underestimates were equally common. Many parameters could be examined, but I chose several important results to evaluate. I examined the time series of total biomass (age 2+) including estimated biomass in the last year (*LY*) and the ratio of last year's biomass to first year's biomass (*LY/FY*) which are important indicators of stock condition. I also examined catchability (q), recruitment variability (σ_r), natural mortality (M), and $F_{40\%}$ (the harvest rate in which spawning biomass per recruit is reduced to 40% of the pristine level, an important proxy for F_{MSY} used in management of groundfish in Alaska. In experiment 3, I also examine the change in estimated Hessian standard deviations and coefficients of variation from the biomass estimates to assess the utility of another index in reducing uncertainty predicted by the model.

5.3 Results

Experiment 1 examined the effect of survey errors on model outputs. My expectation was that measurement errors would not produce substantial biases, but would decrease precision as measurement error increased. The first result that needed to be examined was the base case result where measurement error was zero and survey precision was assumed by the analyst at CV=1%. The base case accurately predicts the biomass parameters and had slight biases in the other parameters (Table 5.3). The results of the base case need to be considered when examining the remainder of the results of this study, in order to be certain that the introduced factors are the cause of the effects. When I compared the results of the first experiment using the first 100 and the second 100 simulations for each experiment, the results were nearly identical which confirmed that 100 simulations per factor was sufficient for the remainder of the study.

The results of the simulations in Experiment 1 showed that the last year biomass (*LY*) was negligibly biased and was relatively precise until the measurement error reached 50% (Tables 5.3-5.4), where the model became very imprecise. At this highest level of measurement error, the biomass estimate became negatively biased by about 15% (Figure 5.1). For most parameters, the estimation precision decreased as the error level increased with a major decrease at the highest measurement error. Catchability (q) was highly negatively biased at the highest measurement error (Figure 5.2). Natural mortality (M), harvest rate ($F_{40\%}$) and recruitment variation (σ_r) were only slightly biased at all measurement error levels (Figure 5.2). The results were as expected, showing that the model performs ineffectively at high levels of survey measurement error. .

In experiment 2, I examined the effects of a time trend in survey catchability. My expectation was that trends in catchability would lead to large changes in precision and accuracy in biomass estimates and that high measurement error would increase the effect. The results were generally consistent with this expectation. At low measurement error, the biomass estimates were precise but much more biased (Figure 5.3), since the catchability trend was clear in the data. At high measurement error, the biomass estimates were imprecise, but much less biased (Tables 5.5-5.6). The precision and accuracy of natural mortality were relatively stable with a small decrease in precision at the highest measurement error. The precision and accuracy of catchability were also relatively stable until the highest level of measurement error. At high measurement error, catchability (q) was underestimated under both positive and negative trends, and at low measurement error, the catchability was overestimated when catchability was decreasing and underestimated when catchability was increasing. At moderate measurement error (25%), harvest rate ($F_{40\%}$) and recruitment variation (σ_r) were stable and exhibited only minor changes with catchability trends (Figure 5.4).

In experiment 3, I examined the effects of adding an additional biomass index to the stock assessment model. My expectations were that adding an additional index that confirmed the trend of the trawl survey index would result in precise estimation, low bias and a lower standard deviation of the biomass estimates than the one-survey model. If the secondary index had a contradictory trend in either direction, the model would produce less precision and a bias in the direction of the trend with an increased standard deviation on the biomass estimates compared to the one-survey model. These expectations were mostly confirmed. Biomass estimates using the index with the confirmatory trend were negligibly biased with higher precision than the one-survey model at each corresponding level of measurement error (Tables 5.3, 5.7, 5.8). The biomass estimates for the secondary indices with downward trends resulted in the lowest precision and were the most biased. The biomass estimates for the secondary indices with upward trends were precise and had small bias (Figure 5.5). Catchability was highly negatively biased by the additional index with no trend. Harvest rate, natural mortality, and recruitment variation were precise and accurate at all factor levels (Figure 5.6). The estimated standard deviations of the biomass estimates (from the Hessian matrix calculated by AD Model builder) were converted to coefficients of variation. When these CVs were examined (Table 5.9, Figure 5.7), they were lower for the increasing trend and higher for the decreasing trend, but the CVs were generally worse as measurement error increased. However when comparing the CV of the generating model (27%) with the CV of the 2-index model, there were only substantial gains (6%) in the precision of the simulations with the lowest measurement error and the corroborating trend.

In experiment 4, I examined the effects of using different weights on the multiple data sources. I expected that the data sets that contribute the most to the objective function total would be most sensitive to changes in parameter weightings. In this experiment and experiment 5, I only examined relative median error because there was only one simulation per configuration. Contrary to my expectation, the only data weighting that had a consistent impact on both biomass and parameter estimates was fishery length data (Table 5.10). Down-weighting this parameter to 10% of its original weight increased ending year biomass almost two-fold (Figure 5.8), while increasing it to 10 times its weight produced a slightly negatively biased estimate. Changing the weight on fishery lengths also had a strong effect on estimation precision of other parameters (Figure 5.9). The remainder of the weightings on other data sources produced minimal biases with a slightly larger increase in bias for lowering weights than raising weights.

In Experiment 5, I examined the sensitivity of results to changes in the prior distributions for key parameters. My expectation was that the model would be most sensitive to the prior distribution of natural mortality, which is very difficult to estimate within stock assessments because it is confounded with many other parameters in the model such as catchability and fishing mortality (Fu and Quinn 2000). Indeed, natural mortality was very sensitive, even though its prior CV was much lower than the other two parameters (Table 5.11). When the mean of the prior was 50% less, it caused a maximum 40% negative bias in the biomass estimates, but when it was doubled it resulted in between a 100 and 250% positive bias in the biomass estimates (Figure 5.10). The relative errors in natural mortality and harvest rate caused by the change in its prior were in the same direction as the biomass, but not as extreme and more symmetrical, while the relative error of catchability was roughly inversely proportional. Recruitment variation was also affected noticeably, but not extremely. Varying the CV of the natural mortality prior produced considerable impacts to the other parameter estimates (Figure 5.11). Changing the prior on catchability had the next largest effect. As the mean of catchability was increased, the ending year biomass estimate decreased from a greater than 76% positive bias to a 12% negative bias when it was doubled at the most precise prior distribution (Figure 5.10). Interestingly, when the catchability prior was loosely specified (CV=40%), the estimates of biomass were relatively unbiased for all values of the prior mean. The changes in the priors on catchability did not have the same inverse effect on natural mortality, which remained fairly consistent for all levels of the catchability prior. Changes in the catchability prior had a minimal effect on relative median error of harvest rate or recruitment variation. The prior on recruitment variation had the least effect (Figure 5.10), with a tendency to overestimate biomass and underestimate catchability at low levels and the opposite effect at high prior means. Figure 5.12 shows how correlated these three key parameters are when they move toward extreme values, with a very small surface where they can take on a variety of values.

5.4 Discussion

These simulation results indicated that modern stock assessment methods, which estimate many parameters and incorporate multiple data sources, perform in complex ways. The model is a way of extracting information on stock size and viability by tracking the age of fish and the lengths of fish through time. These data are tuned to the right order of magnitude by the catch data and the survey biomass index. This combination is then used to estimate recruitment and spawners in the population to give information about current and future biomass so that the fishery can be managed optimally.

The part that survey error plays in the stock assessment model is one of the more common themes in fisheries management. This theme involves tradeoffs between costs of additional survey effort and whether that additional effort results in diminishing returns in stock assessment precision. Aggregated species such as rockfish exhibit this concern more than more uniformly distributed species due to high variances associated with any high biomass estimates (Hanselman et al. 2001, 2003b). In experiment 1, I simulated the effect of measurement error in the survey biomass estimates on the results from the stock assessment model. The results of the simulations showed that measurement error had a substantial effect only at high measurement error (CV=50%). There was a negative bias in biomass estimation at the highest measurement error level, which was inconsistent with other studies that used the same prediction model as the generating model (e.g. Bence et al. 1993, Sampson and Yin 1998). The catchability bias was constantly negative, but small. This showed that the model did not confuse measurement error with a change in catchability until the index became highly variable. The introduction of error into the biomass index did not have much effect on the beginning of the time series because the bias of the ratio of last year to start year results were quite similar to the bias of the last year biomass. Other parameters were negligibly affected. In general, when measurement error was low, there was little effect on the stock assessment results, whereas when it was high, the stock assessment model did not perform effectively.

Another factor that has implications in the stock assessment process is the role of a trend in catchability of the species of interest. Model misspecification due to incomplete knowledge often results in large relative errors in population estimates (NRC 1998). Trends in catchability can cause a much different estimate of biomass in the current year than in the future as more data are added (Parma 1993). Positive trends could be caused by survey teams becoming more efficient with the standard gear and the addition of more advanced net sensor data to ensure the net is fishing properly, or more fish moving into trawlable areas because of changes in environmental conditions. Negative trends could be caused by more of the prime rockfish habitat being designated as untrawlable by the surveys due to net damage, or due to rising sea temperatures increasing the ability of fish to swim out of the net. In the simulations in experiment 2, I showed the effects of upward and downward trends in catchability over the survey period in the model (8 surveys, 1984-2003). At high measurement error, trends in catchability acted as expected by producing higher estimation CVs, but biases were most substantial at low measurement error. This was because at low measurement error, the trend was clear in the data and the lower CV gave the biomass estimates more weight in the estimation. When the measurement error was highest, the simulations result in a substantial and similar estimation error for both negative and positive catchability trends. The model seemed to have no ability to detect this changing catchability in the survey. When measurement error was high, it tended toward a higher catchability for the increasing catchability trends, but was negatively biased for all trends. When the measurement error was low, the model was completely inept in detecting this trend, in fact estimating catchability in opposite directions to the trend. While the model estimates an average catchability over time, a positive trend would reflect a higher overall survey mean and the model should lower the average catchability accordingly. Other parameters were relatively unaffected by the trends in catchability.

Often in stock assessment it is important to consider adding additional indices of biomass. Recently, there has been interest for rockfish in Alaska to add an additional survey specific to rockfish or to incorporate ship-board hydroacoustic data into a model index. For experiment 3, I simulated the potential of this by adding an additional biomass index for the last five survey years. I expected that when this index confirmed the trawl survey index it would result in better precision and lower bias than when the index was contradictory to the trend of the trawl survey. The model would also perform better than the one-survey model. I also expected that an additional index would lower the estimated standard deviation of total biomass from the Hessian matrix produced by AD Model Builder from the one-survey model, with the largest reduction produced by the confirming index. The confirming trend did result in more precise biomass estimates than the one-survey model at all factors, but were similarly negatively biased at the highest level of measurement error. A second index with a decreasing trend had a larger effect on the model results than an index with an upward trend. The trawl survey's catchability was generally underestimated with the addition of an additional index, whereas the remainder of the parameters were relatively unaffected. Unexpectedly, the coefficient of variation of total biomass calculated from the Hessian matrix only decreased for two of the scenarios, the lowest measurement error cases with no trend and the positive trend. The overall results of Experiment 3 indicated that the model performed better with the second index, but this was not necessarily reflected in estimated standard deviations from AD Model Builder.

How stock assessment practitioners weight different data sources has often been contentious and considered a duplicitous way of getting desired model results. While the data source is usually weighted by the inverse of some function of its variance, it is often up to the stock assessment biologist to put some perception about the quality of that data into the model in terms of a weighting factor (Merritt and Quinn 2000). This is often done implicitly by assuming all data weights to be one, which means that the biologist believes all the data is of equal quality. Or these perceptions can be applied explicitly, by weighting data sources deliberately. In the base model presented here, I use the implicit assumption of all data being equal except for catch, which I give a weight of 50, because I believe it was measured with some degree of precision relative to other data sources. In experiment 4, I tested the sensitivity of these weighting assumptions by adjusting the weights of each data source. In the simulations in Experiment 4, the only data weighting that had a substantial effect was the fishery length data. Fishery lengths and catch were the only data sources in the model that have information prior to 1984. Changing the weight of fishery lengths had a large impact on the estimates of ending year biomass, catchability, and recruitment variation. The larger ending biomass with lower levels of weighting on the fishery lengths was a result of several factors: (1) The model was more influenced by recent data, which showed an upward trend. (2) Catchability estimates were much lower, implying that the trawl survey did not catch more fish than the area swept, as the base model predicted. (3) Recruitment variation was lower, implying a steadier recruitment over time. Therefore, it was not the data source with the highest proportion of the objective function that was the most sensitive, but the one with the most data. An interesting examination of the effect of having one data source that spans a longer time period than other data would be prospectively removing years from the beginning of that time series until all the time series of data start in the same year. The results in Experiment 4 were contrary to those of Radomski et al. (*in press*[13]) who found that weighting terms had significant effects on the error of population estimates.

Perhaps the most controversial part of a Bayesian or semi-Bayesian stock assessment model is the use of prior distributions. Recently, some practitioners have accepted that it is better to employ an informative prior and test its sensitivity, rather than only use uninformative priors (Punt and Hilborn 1997). Meta-analyses such as Myers et al. (1995) have allowed for reasonable informative priors to be constructed. In the base

---

[13] Radomski, P., T.J. Quinn II, and J.R. Bence. *In press.* Comparison of virtual population analysis and statistical kill-at-age analysis for a recreation kill-dominated fishery. Transactions of the American Fishery Society.

model of this simulation, I use reasonably informative priors to estimate catchability and recruitment variability and a highly informative prior for natural mortality. Experiment 5 tests the sensitivity of these priors. The results of these simulations showed that the prior distribution of natural mortality, even though constrained more than the other parameters (because of its difficulty in estimation), had large implications on the model outputs. When the CV of the prior distribution for natural mortality was doubled, the bias of the last year biomass doubled. When the estimate of the prior mean of natural mortality was doubled, last year biomass was overestimated by over 200%. The resulting natural mortality estimates from changing the prior information was roughly inversely proportional to the estimates of catchability. Interestingly, this relationship was only one-way as changes in the catchability prior only affected the estimates of catchability and not natural mortality. Likely, this was due to the tight prior imposed on natural mortality on the base model. The prior distributions on catchability caused some sensitivity in the biomass estimates, but about half as much as natural mortality. When the CV of the prior distribution on catchability was made less precise, biomass estimates were nearly unbiased for all prior means. The prior information on recruitment variability had very little effect across all parameters.

5.5 Management implications

Separable age-structured stock assessment models conducted in AD Model Builder have provided a new level of flexibility for synthesizing multiple data sources, estimating multitudes of parameters, incorporating process error, and using Bayesian techniques. Along with this flexibility is added complexity for the assessment biologist for both deciding what will go into the model and interpreting the results. Therefore, a necessary step in stock assessment is for these models to be explored and validated through simulations.

In this study, I explored just some of the many factors involved in stock assessment. The first experiment I looked at was the effects of survey measurement error in a one-index stock assessment model. These simulations would suggest that measurement error is not a major concern for species with uniform distributions that have reasonable (<25%) levels of measurement error. However, species like rockfish that have high estimated survey error on the order of CV=50% may suffer from some significant biases in their stock assessments. This much measurement error in the survey makes accurate stock assessment improbable. The experiment confirms that improving the survey design or sample size of the survey to lower this CV is worthwhile to improve stock assessment precision.

The second experiment I looked at was the effect of changing catchability over time. These simulations showed that a positive trend in catchability had a stronger impact on stock assessment results than a negative trend. The trends do not have a big impact on starting biomass which can have important implications on management quantities such as virgin biomass B_0 which is used to assess the current status of the stock. Additionally, at high survey CVs, the model could slightly detect that there was a trend in the catchability, while at low survey CVs, the model estimated the trend in the wrong direction. This means that if a trend in catchability is suspected in a survey, it can have important implications on the model results and should be accounted for.

In the third experiment, I showed the effect of adding an additional biomass index. Close attention has been paid to adding additional biomass indices to reduce uncertainty in stock assessments. This experiment showed that an additional survey index increased precision at every measurement error level when compared to the one-survey model in Experiment 1. The increase in precision at the highest measurement error level indicated that stock assessment for rockfish in Alaska would benefit in terms of model estimation reliability with a secondary index. A future experiment of interest would be to omit the original survey index and use only the shorter, more precise index or to determine how long the time series must be before the inferior index is omitted.

In the fourth experiment, I analyzed the effect of data weights in the model specification process. The simulation showed that the only data source weighting in this model that made a large difference was the weighting on fishery length data. This was interesting in that many models use historical length data of unknown quality that sets the stage for future biomass and recruitment. Stock assessment scientists should take a close look at historic length data they are using and its effect on assessment results.

Finally, in the fifth experiment, I examined the sensitivity of the model to prior distributions used to estimate key parameters. Overwhelmingly, model results were most sensitive to the tight prior distribution for natural mortality. Misspecification of natural mortality can have enormous repercussions on biomass estimates and harvest rates[12]. Obtaining reliable independent estimates of this parameter would be very useful, but are extremely difficult to ascertain for rockfish. Furthermore, natural mortality was the parameter with the greatest uncertainty, but the most precise prior distribution is used for it to force the model to converge to reasonable results. This inherently results in a facade that the results are more certain than they really are. Catchability was not as sensitive and it appears that setting it at a reasonable value with a loose prior CV is adequate, but a comprehensive study on its actual value is needed. Recruitment variation was relatively insensitive to its prior distribution.

This study represents only five aspects that could be investigated on these types of models. Future studies of these types of models should include: (1) Assessment of aging error and aging sample sizes (Coggins and Quinn 1998), (2) Evaluation of error in maturity and growth schedules, (3) Assessment of different selectivity curves such as dome-shaped and asymptotic, and time-varying selectivities[13]. Aside from these possibilities, there are many avenues that must be explored to continue validating the results of these stock assessment models.

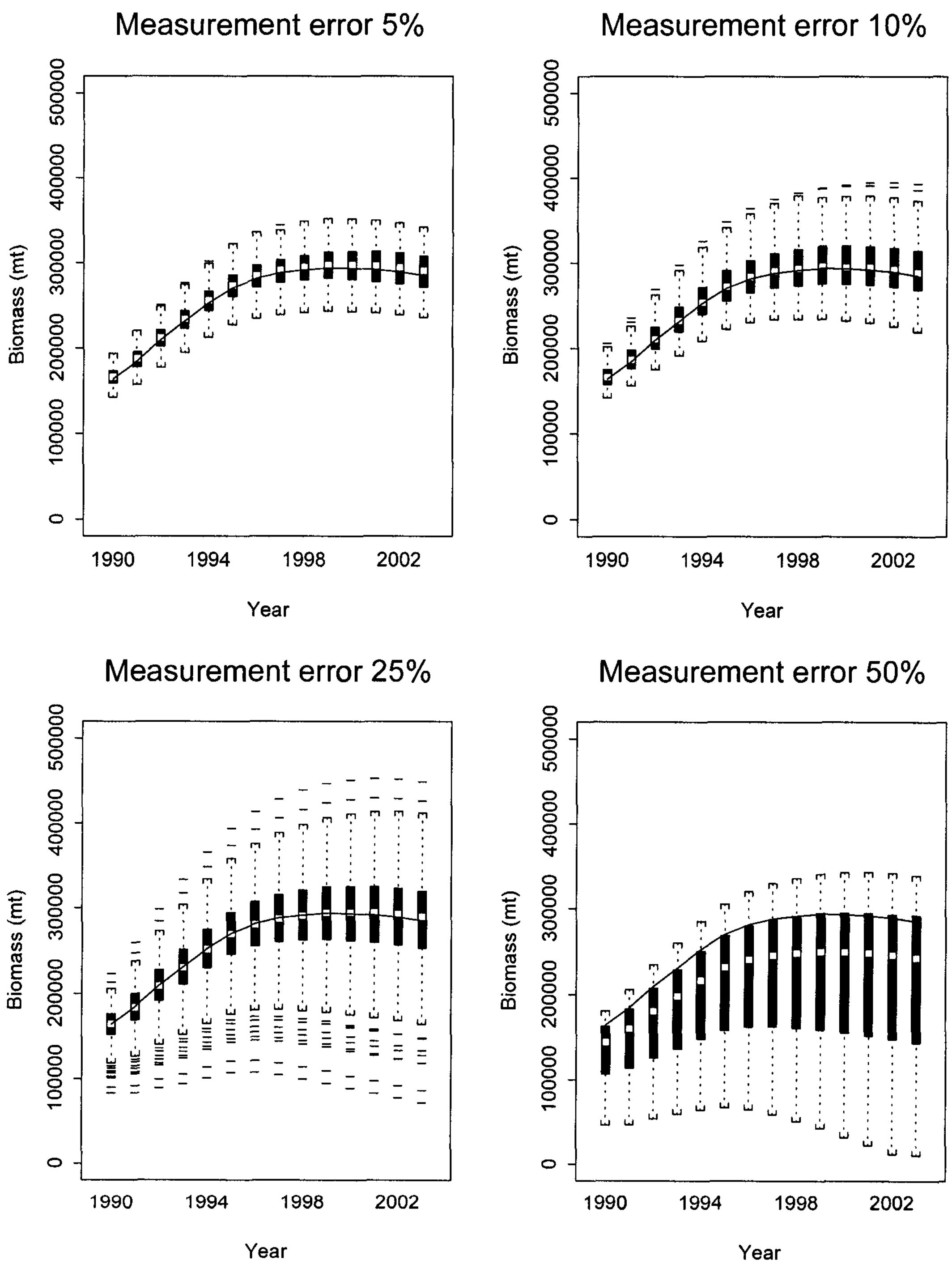


Figure 5.1. Effects of measurement error on total biomass from 1990-2003 for Experiment 1. Solid line is "true" total biomass.

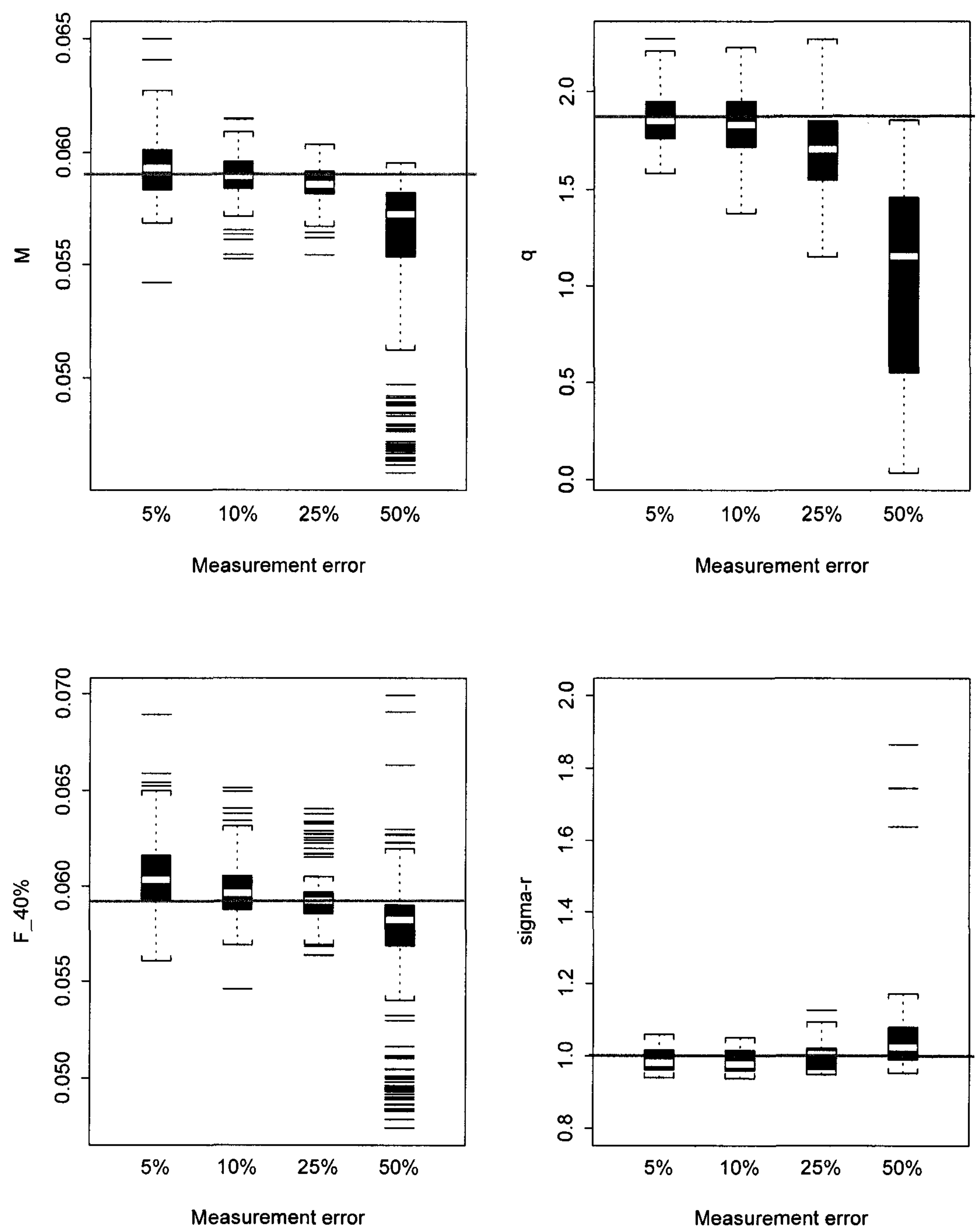


Figure 5.2. Effect of measurement error on key parameter estimates for experiment 1. Solid line is "true value".

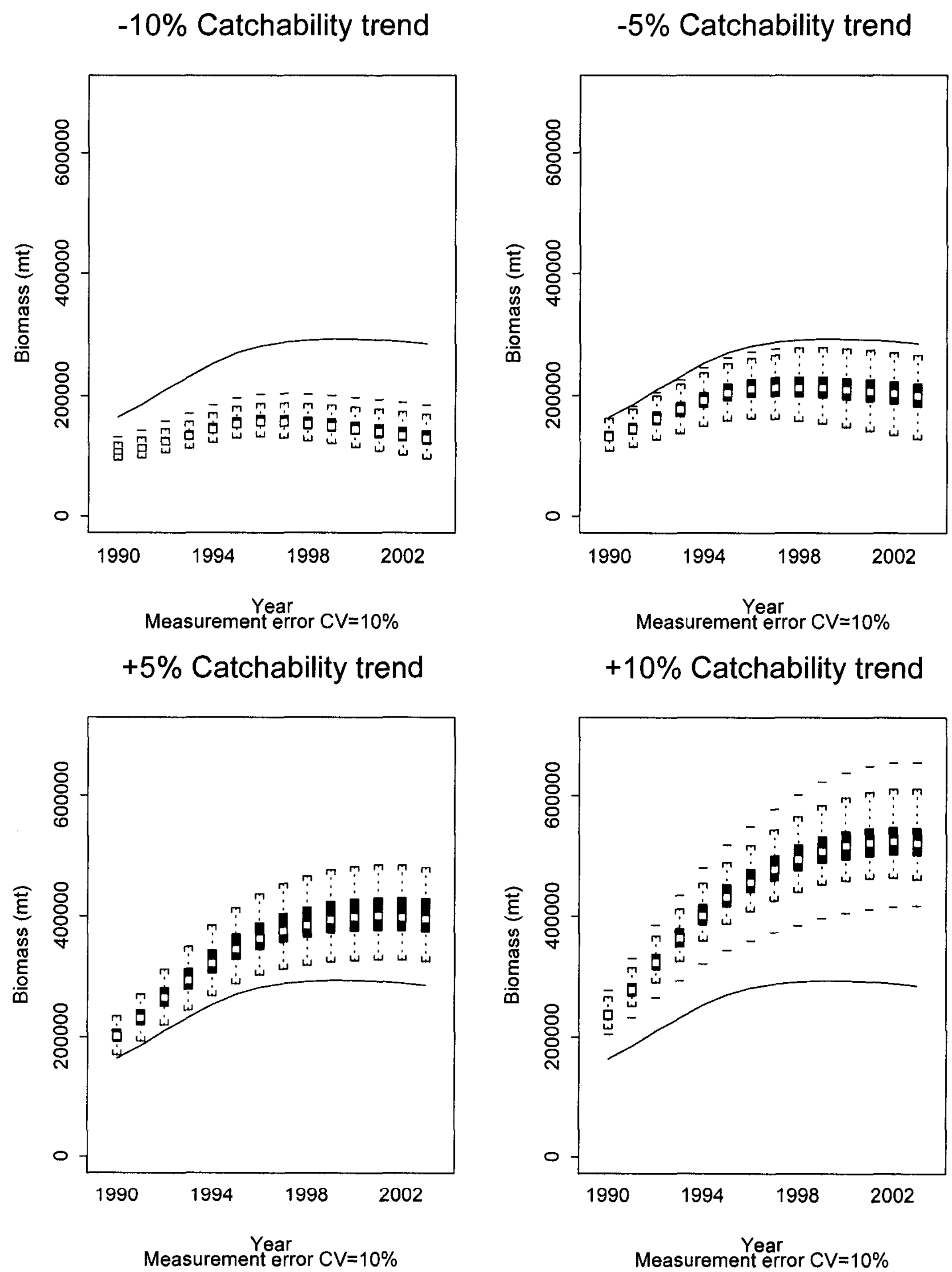


Figure 5.3. Boxplots of distribution of total biomass from 1990-2003 for experiment 2 examining catchability trends. Solid line is "true".

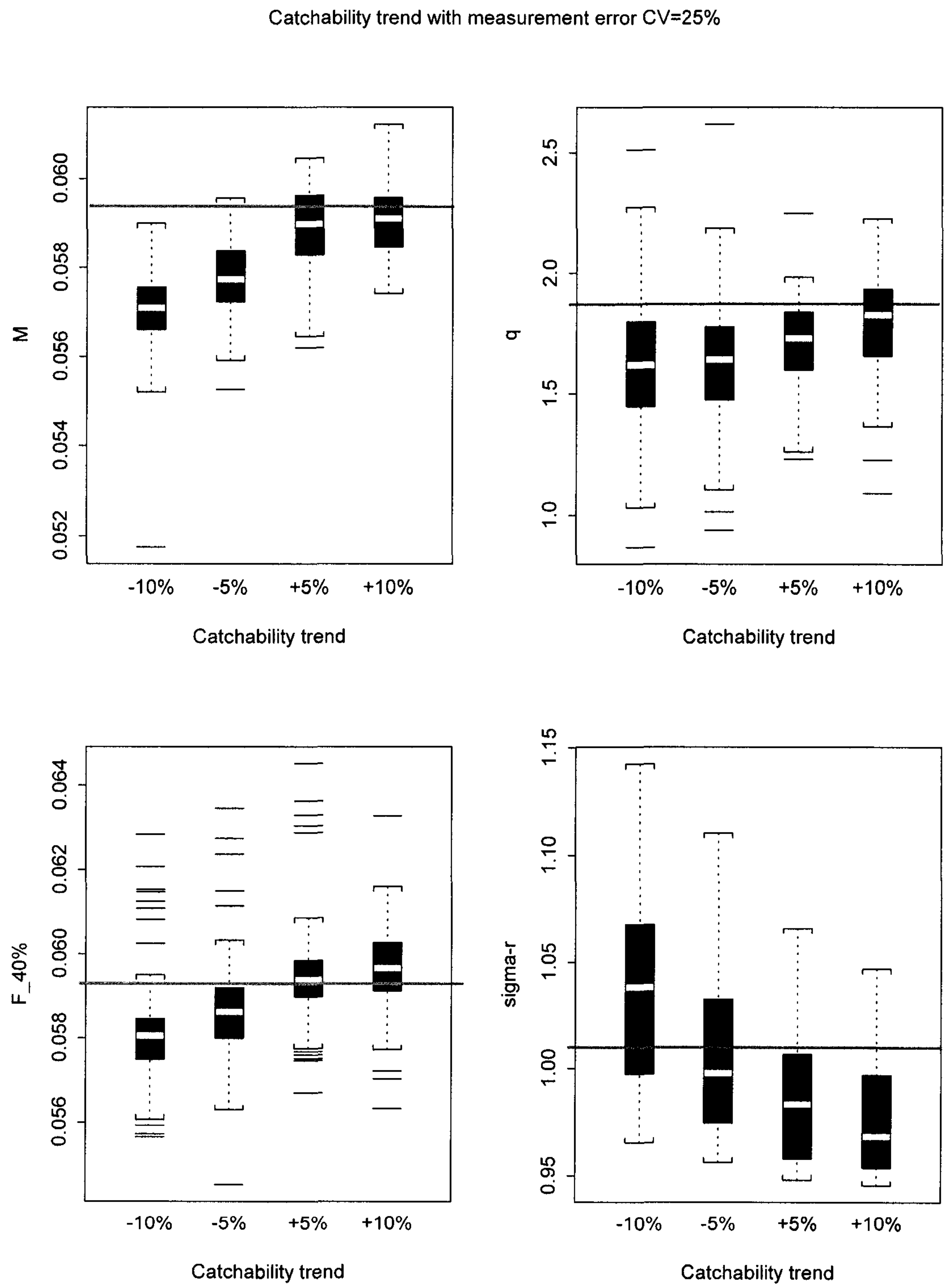


Figure 5.4. Boxplots of distribution of key parameters for Experiment 2. Solid line is true.

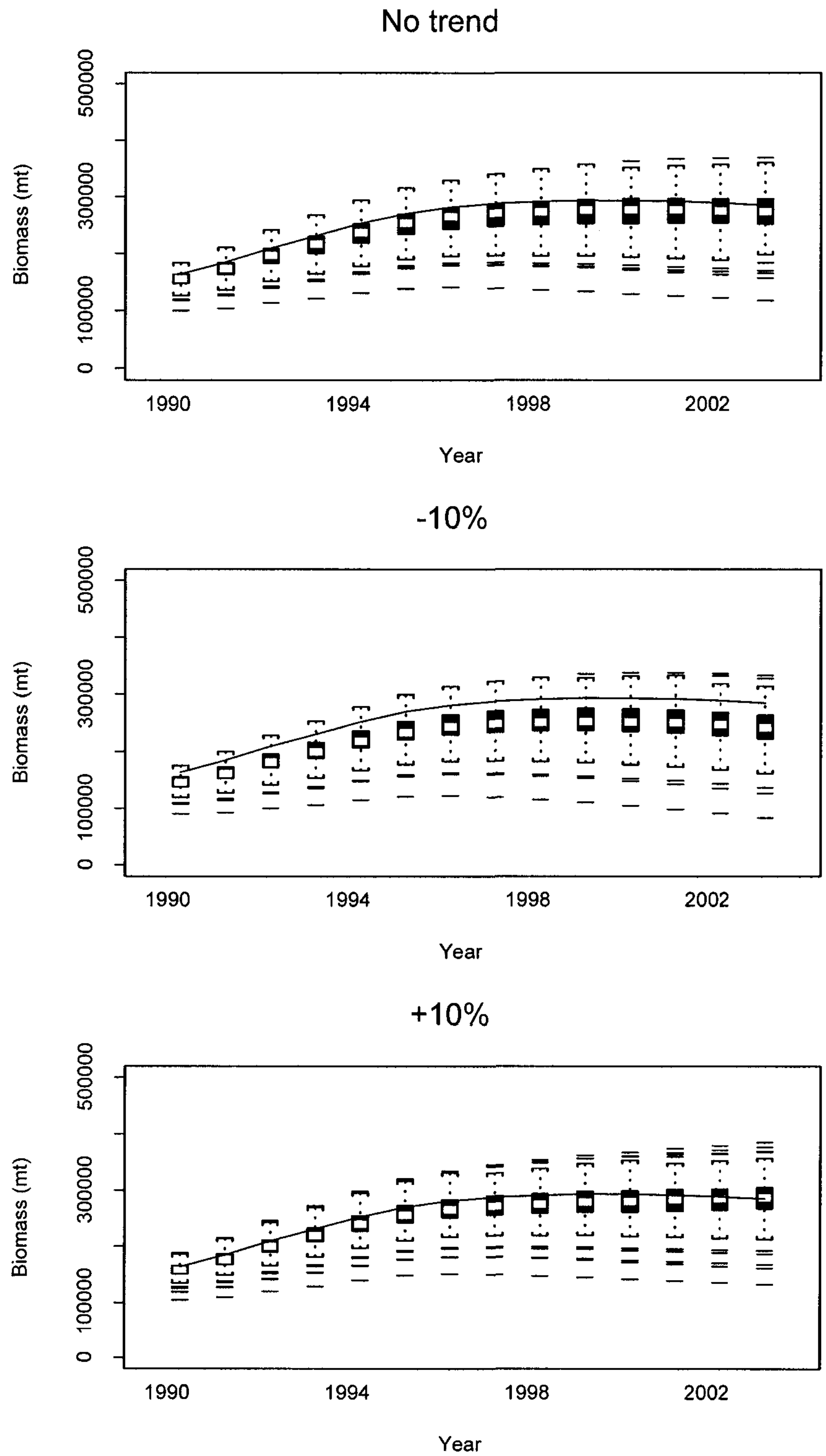


Figure 5.5. Boxplot distributions of total biomass from 1990-2003 with a secondary biomass index with varying trends for Experiment 3 with 25% and 10% measurement error on Indices 1 and 2, respectively. Solid line=true biomass.

Parameters with measurement error for indices of 25/10%

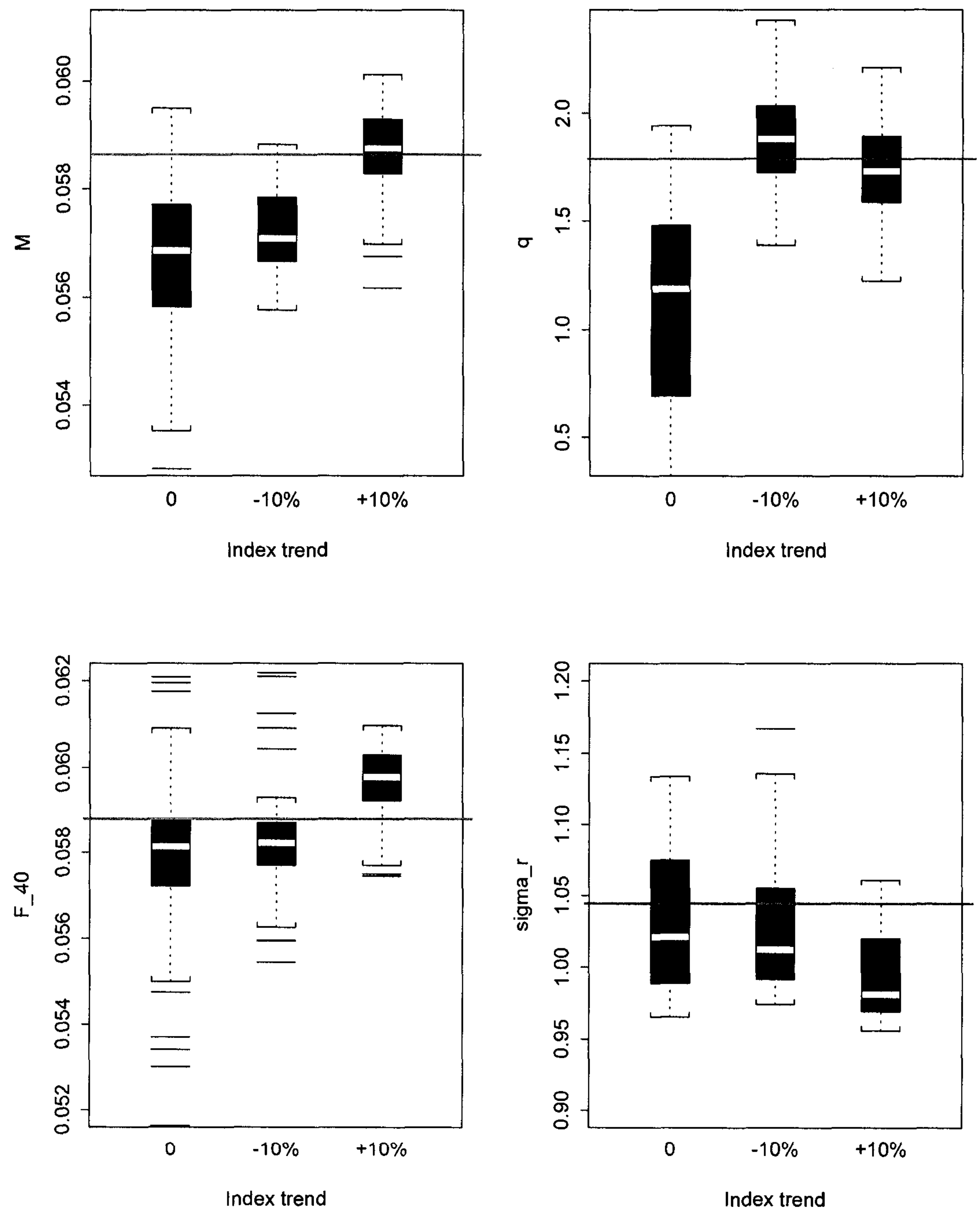


Figure 5.6. Boxplot distributions of key parameters for Experiment 3 with 25% and 10% measurement error on Indices 1 and 2, respectively. Solid line=True.

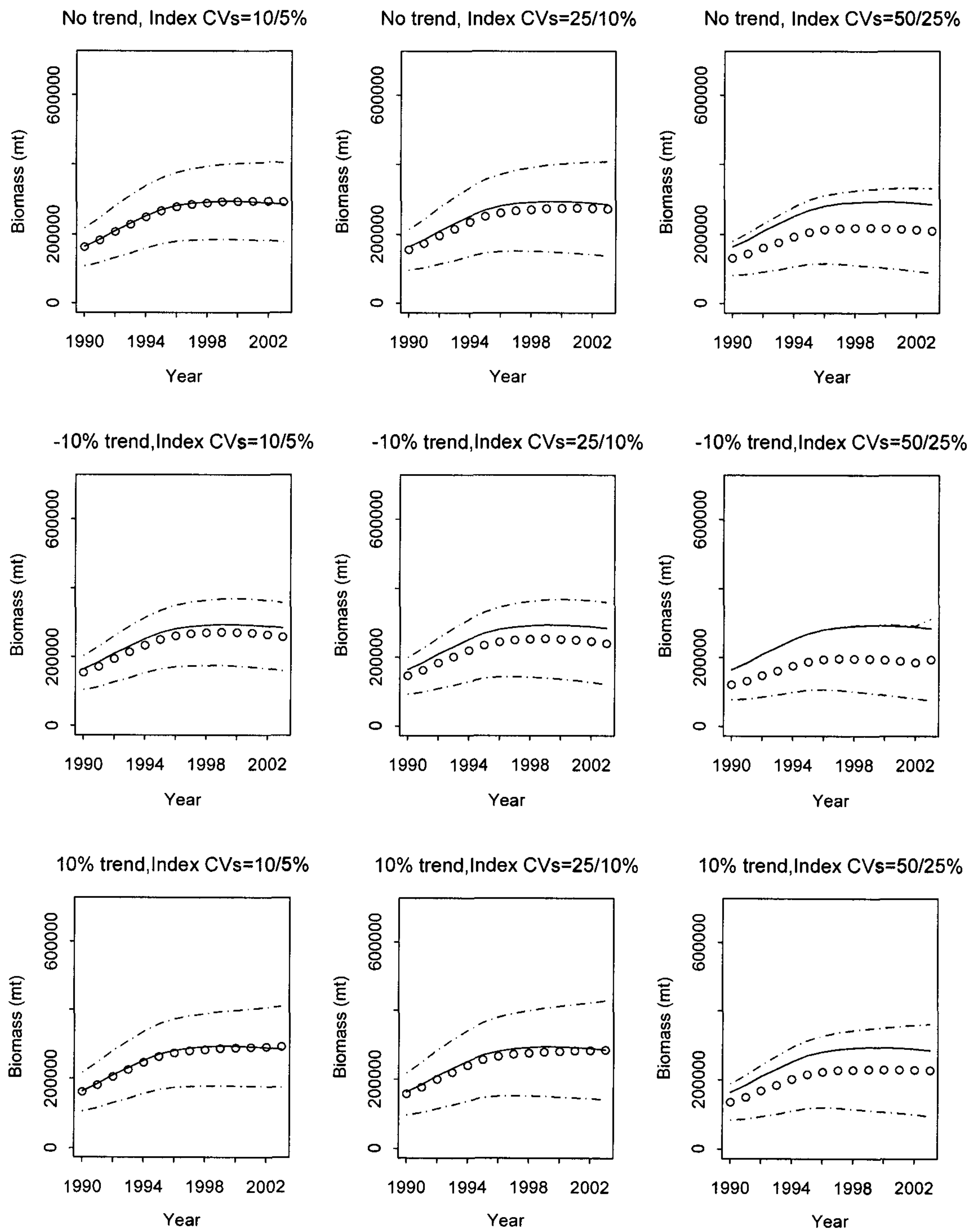


Figure 5.7. Effect of adding a secondary index with various trends with 25% and 10% measurement error for the first and second indices, respectively, on total biomass from 1990-2003 for experiment 3, with 95% confidence intervals (dashed lines) from AD Model Builder Hessian matrix. Circles=predicted, solid=true.

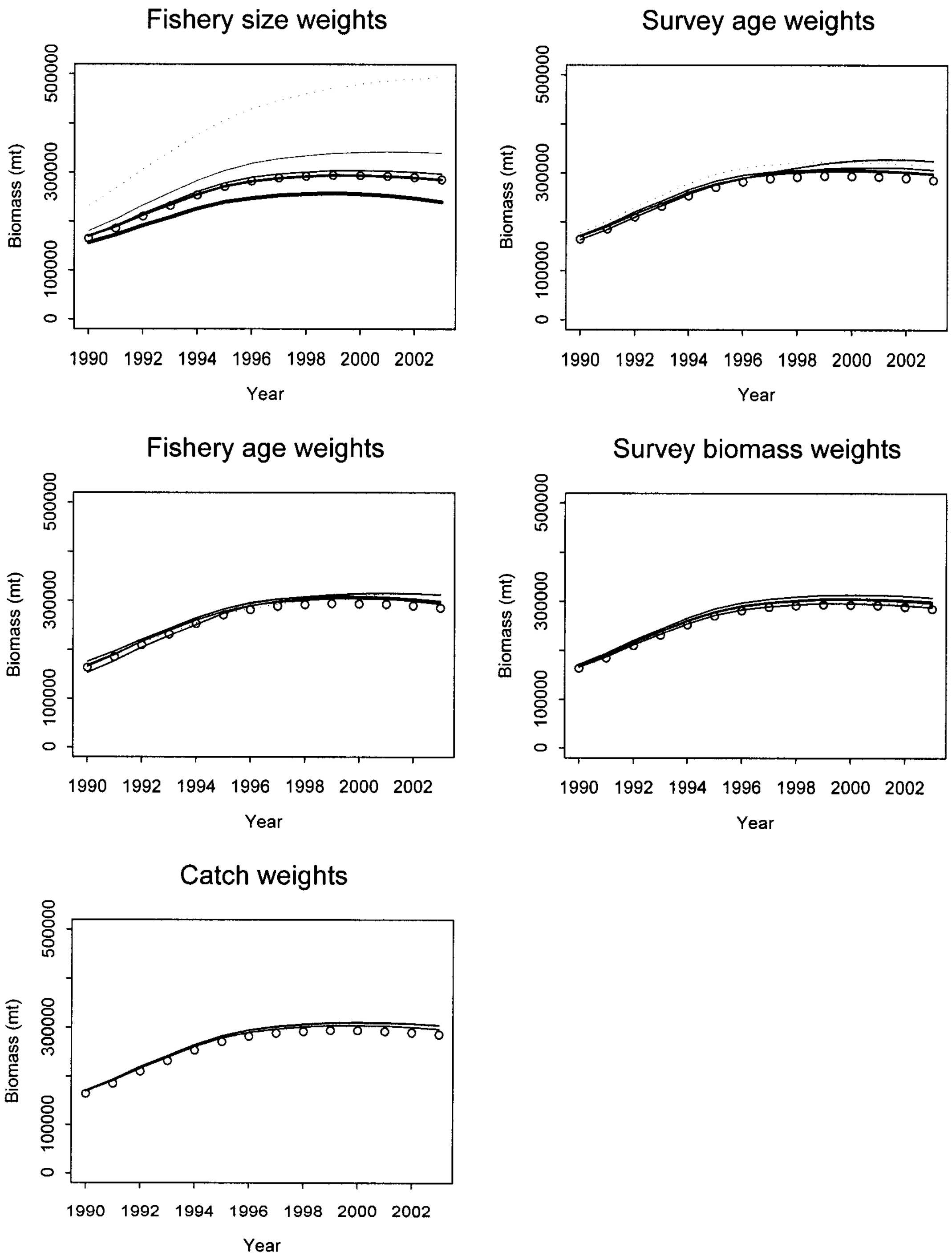


Figure 5.8. Biomass trajectories for different weighting factors for Experiment 4, dotted line is least weight, and dark line is most weight

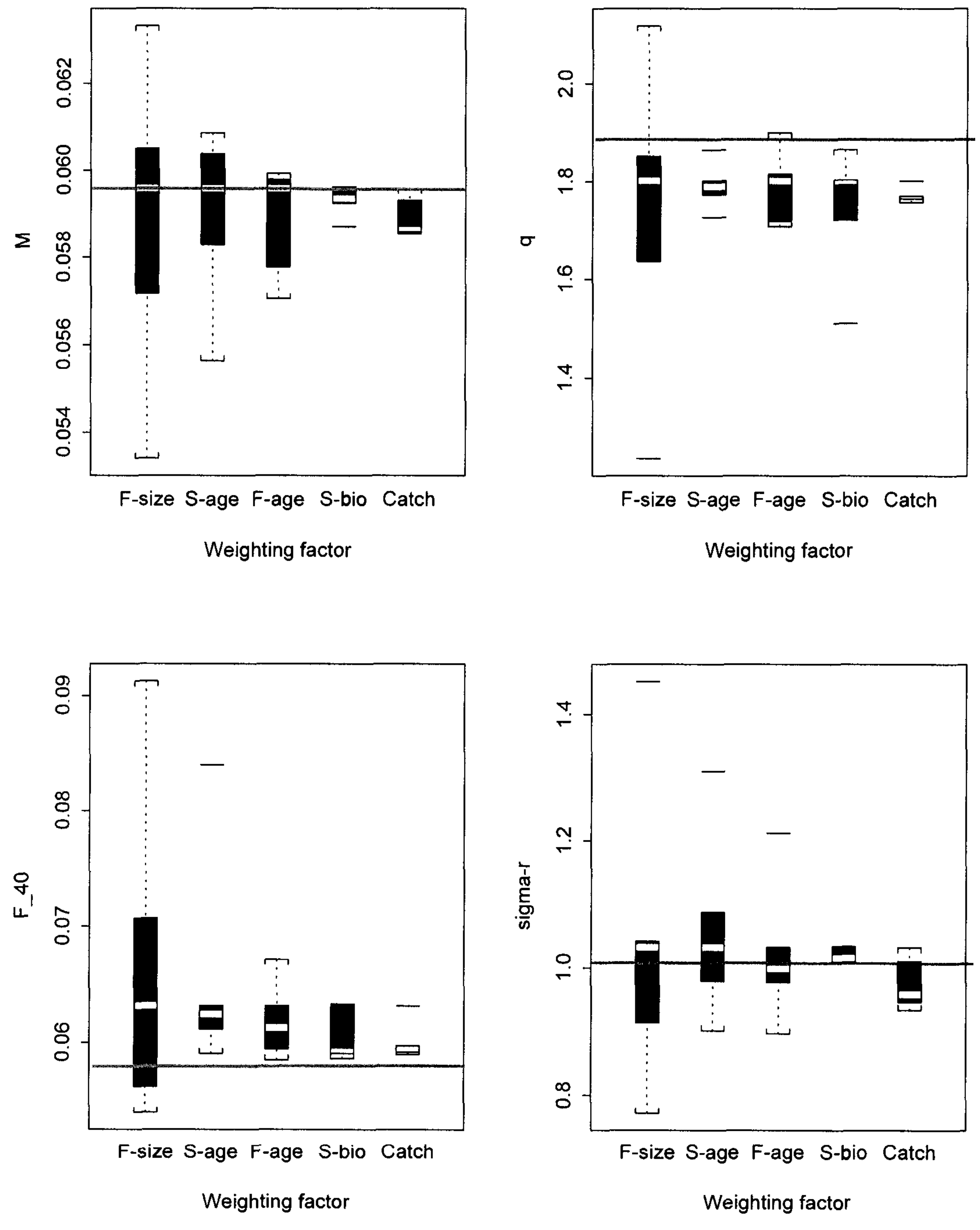


Figure 5.9. Boxplot distributions of key parameters when different data sources' weights are varied for experiment 4. Solid line=True.

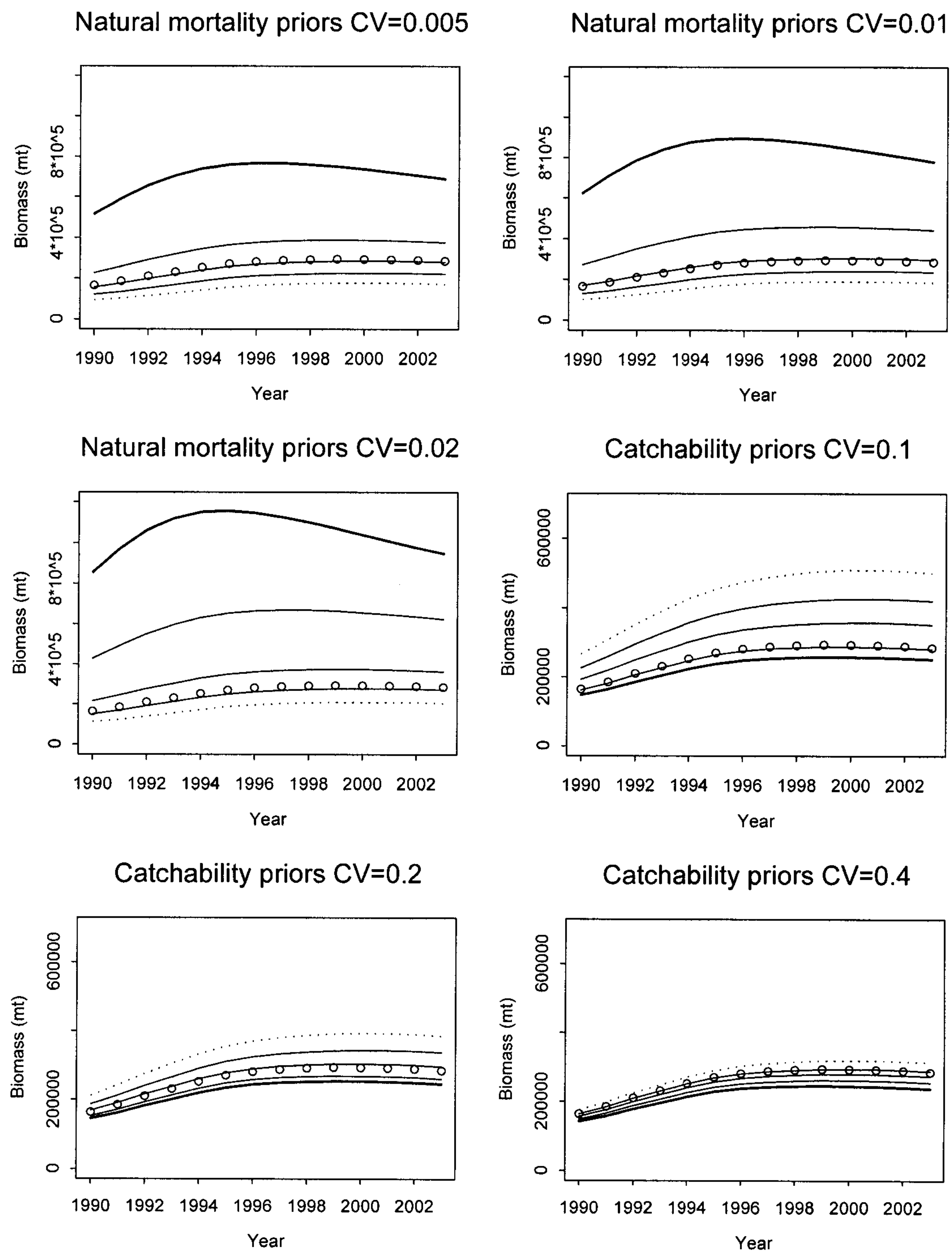


Figure 5.10. Biomass trajectories for different prior CVs for Experiment 5, dotted line is lowest prior mean, and dark line is highest prior mean.

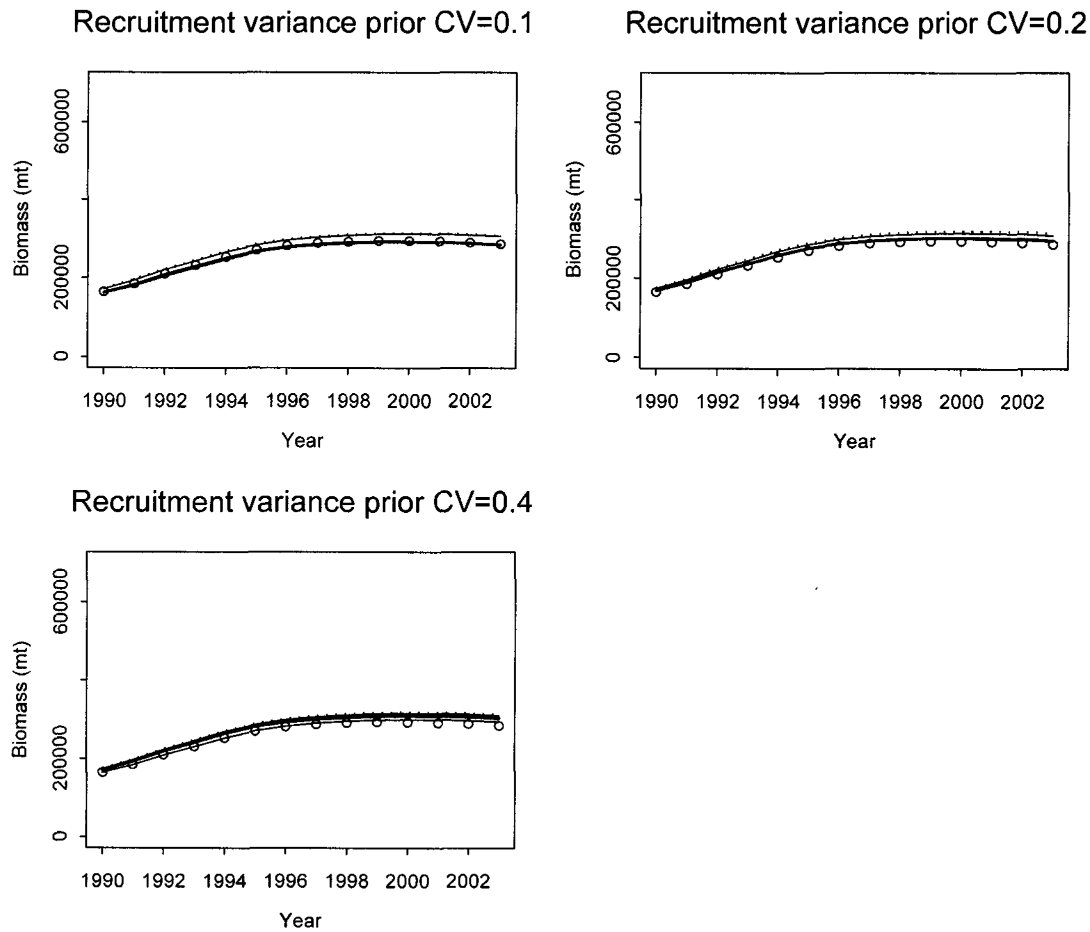


Figure 5.10 (cont.). Biomass trajectories for different prior CVs for Experiment 5, dotted line is lowest prior mean, and dark line is highest prior mean.

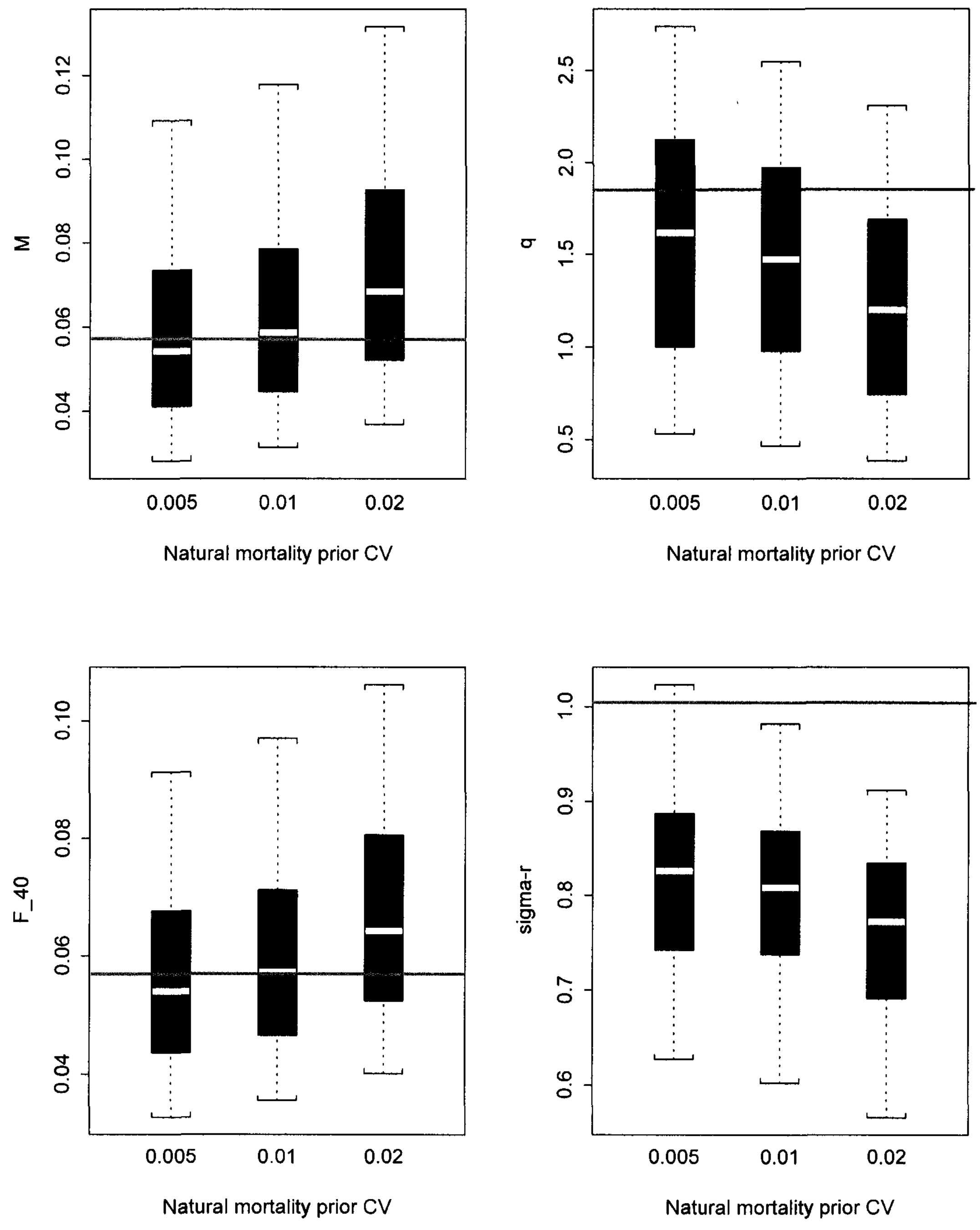


Figure 5.11. Boxplot distributions of key parameters as the CV of the natural mortality prior is varied in experiment 5. Solid line=True.

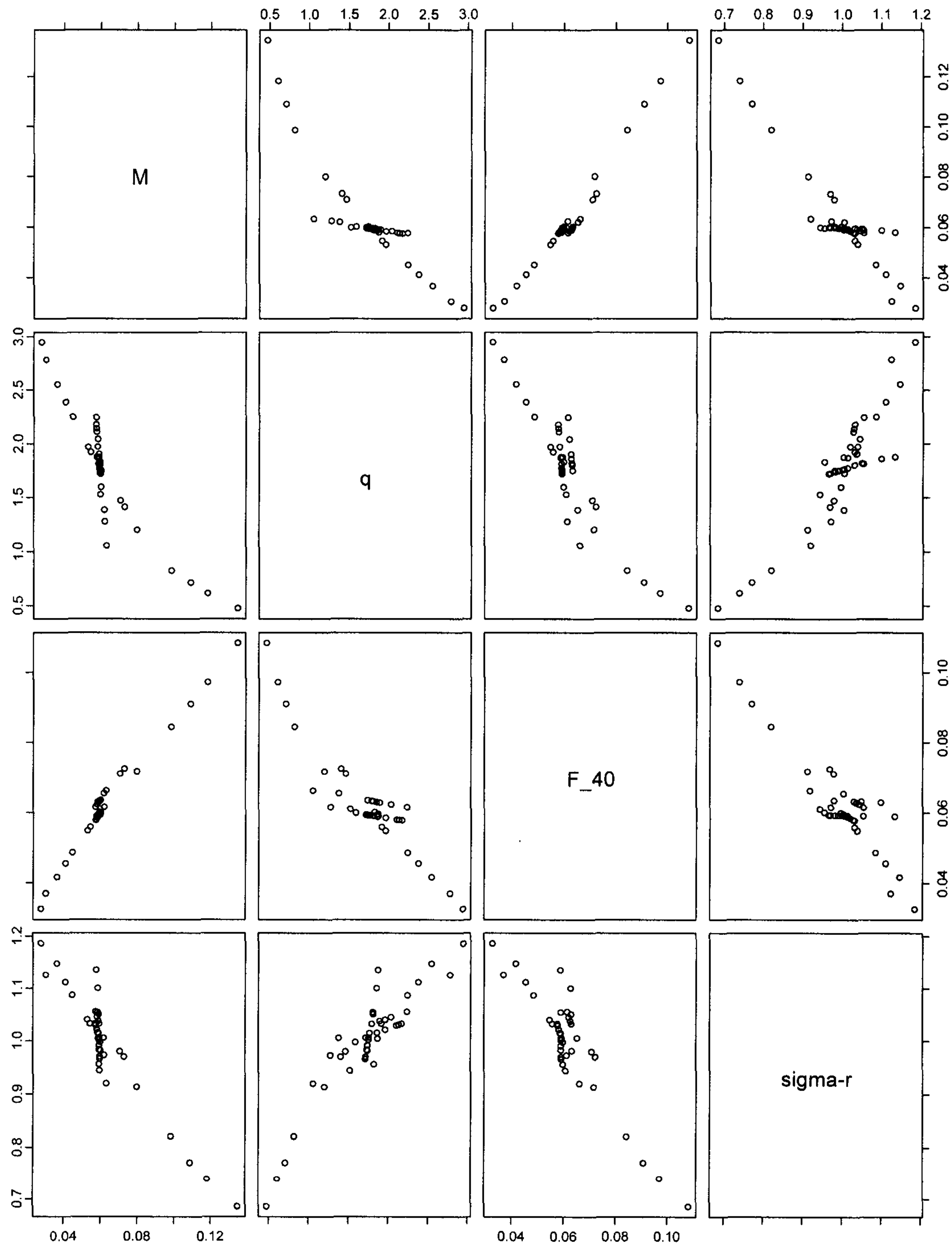


Figure 5.12. Scatterplot demonstrating the high correlation between key parameter values in Experiment 5.

Table 5.1. Data sources used in Pacific ocean perch generating model

Data Type	Number of years	Total data points
Catch	43 (1961-2003)	43
Survey Biomass	8 (1984-2003)	8
Survey Ages	6(1984-1999), 24 ages	144
Fishery Ages	5(1998-2002), 24 ages	120
Fishery Lengths	21(1963-1977, 1990-1992, 1995-1997), 23 lengths	483

Table 5.2. Summary of experiments.

Experiment	Effects	Treatments	Runs
Base Case	No measurement error, estimated survey precision CV = 1%	None	1
1	Survey Measurement Error CV (4)	5, 10, 25, 50%	800
2	Catchability trend (4)	-5,-10,+5,+10%	1600
	Measurement error CV (4)	5, 10, 25, 50%	
3	Second survey index trend (3)	0, -10, +10%	900
	Measurement error CV combinations of Index 1 to Index 2	10/5%, 25/10%, 50/25%	
4	Data source weightings (5)	10, 50, 100, 500, 1000%	25
5	Alternate prior means (5)	50, 75, 100, 133, 200%	45
	Alternate prior CVs (3)	50, 100, 200%	

Table 5.3. Experiment 1 relative estimation precision (REP) measured in coefficient of variation (CV) for examining survey measurement error. *LY*=last year biomass, *LY/FY*=last year biomass/first year biomass, *M*=natural mortality, *q*=catchability, $F_{40\%}$=harvest rate, σ_r=recruitment variation.

	Measurement error CV			
Parameter	5%	10%	25%	50%
LY	8.3%	11.5%	21.6%	52.1%
LY/FY	7.8%	11.0%	21.0%	51.0%
M	2.3%	1.7%	1.4%	7.3%
q	7.3%	8.9%	13.6%	58.8%
$F_{40\%}$	3.1%	2.5%	2.2%	6.3%
σ_r	3.1%	3.1%	3.5%	50.4%

Table 5.4. Experiment 1 relative estimation accuracy (REA) measured in relative median error examining survey measurement error. *LY*=last year biomass, *LY/FY*=last year biomass/first year biomass, *M*=natural mortality, *q*=catchability, $F_{40\%}$=harvest rate, σ_r=recruitment variation.

	Measurement error CV				
Parameter	0%	5%	10%	25%	50%
LY	0.0%	1.9%	1.3%	1.7%	-15.2%
LY/FY	0.0%	1.7%	1.5%	1.7%	-14.3%
M	-1.1%	0.3%	-0.3%	-0.9%	-3.5%
q	-1.0%	-1.3%	-2.5%	-9.2%	-41.7%
$F_{40\%}$	-0.3%	0.8%	-0.3%	-1.0%	-3.0%
σ_r	-3.5%	-3.7%	-4.1%	-1.5%	0.5%

Table 5.5. Experiment 2 relative estimation precision (REP) measured in coefficient of variation (CV) examining catchability trends. *LY*=last year biomass, *LY/FY*=last year biomass/first year biomass, *M*=natural mortality, *q*=catchability, $F_{40\%}$=harvest rate, σ_r=recruitment variation.

	Catchability Trend			
Measurement error CV	-10%	-5%	5%	10%
LY				
CV=0.05	8.3%	7.9%	6.3%	5.4%
CV=0.10	12.8%	14.1%	8.6%	7.0%
CV=0.25	21.5%	22.3%	18.2%	12.6%
CV=0.50	50.9%	39.1%	53.0%	51.0%
LY/FY				
CV=0.05	7.7%	7.6%	6.1%	5.4%
CV=0.10	12.4%	13.5%	8.3%	7.0%
CV=0.25	20.9%	21.5%	17.7%	12.3%
CV=0.50	49.7%	38.1%	51.8%	49.8%
M				
CV=0.05	2.8%	2.3%	2.0%	1.9%
CV=0.10	2.1%	1.8%	2.1%	1.6%
CV=0.25	1.6%	1.5%	1.5%	1.4%
CV=0.50	7.0%	6.4%	7.9%	8.4%
q				
CV=0.05	5.5%	5.9%	6.5%	5.5%
CV=0.10	8.4%	10.0%	7.6%	5.7%
CV=0.25	16.9%	16.3%	11.0%	11.0%
CV=0.50	57.0%	47.8%	58.2%	59.0%
$F_{40\%}$				
CV=0.05	3.0%	3.6%	3.3%	2.5%
CV=0.10	3.2%	2.9%	2.5%	2.2%
CV=0.25	2.2%	2.1%	2.0%	1.7%
CV=0.50	6.3%	5.3%	6.8%	6.5%
σ_r				
CV=0.05	3.1%	3.6%	3.0%	2.0%
CV=0.10	3.3%	3.4%	2.8%	2.3%
CV=0.25	4.1%	3.4%	2.8%	2.5%
CV=0.50	51.6%	49.1%	52.9%	55.8%

Table 5.6. Experiment 2 relative estimation accuracy (REA) measured in relative median error examining catchability trends. *LY*=last year biomass, *LY/FY*=last year biomass/first year biomass, *M*=natural mortality, *q*=catchability, $F_{40\%}$=harvest rate, σ_r=recruitment variation.

	Catchability Trend			
Measurement error CV	-10%	-5%	5%	10%
LY				
CV=0.05	-58.0%	-34.1%	47.4%	98.2%
CV=0.10	-54.8%	-30.3%	38.3%	82.5%
CV=0.25	-40.2%	-18.5%	23.8%	44.3%
CV=0.50	-33.3%	-20.3%	-6.5%	1.1%
LY/FY				
CV=0.05	-56.7%	-33.3%	45.8%	95.6%
CV=0.10	-53.4%	-29.3%	36.7%	79.5%
CV=0.25	-38.9%	-17.6%	22.8%	42.1%
CV=0.50	-31.9%	-19.2%	-5.9%	1.5%
M				
CV=0.05	-2.7%	-0.7%	-1.9%	-4.3%
CV=0.10	-3.3%	-1.7%	-0.7%	-2.1%
CV=0.25	-3.5%	-2.4%	-0.3%	-0.1%
CV=0.50	-4.2%	-3.7%	-2.2%	-2.4%
q				
CV=0.05	19.5%	9.6%	-6.1%	-6.3%
CV=0.10	11.1%	5.5%	-5.4%	-6.1%
CV=0.25	-13.9%	-12.7%	-8.0%	-2.8%
CV=0.50	-46.5%	-41.6%	-30.4%	-32.4%
$F_{40\%}$				
CV=0.05	-0.8%	1.0%	-2.1%	-4.8%
CV=0.10	-2.8%	-1.1%	-1.0%	-2.6%
CV=0.25	-3.0%	-2.0%	-0.8%	-0.3%
CV=0.50	-3.9%	-3.4%	-2.3%	-2.2%
σ_r				
CV=0.05	2.7%	-0.3%	-4.9%	-3.0%
CV=0.10	4.6%	-1.1%	-5.2%	-3.9%
CV=0.25	1.8%	-2.1%	-3.6%	-5.0%
CV=0.50	2.8%	0.7%	-0.6%	-1.9%

Table 5.7. Experiment 3 relative estimation precision (REP) measured in coefficient of variation (CV) examining a second biomass index. *LY*=last year biomass, *LY/FY*=last year biomass/first year biomass, *M*=natural mortality, *q*=catchability, $F_{40\%}$=harvest rate, σ_r=recruitment variation. CVs are in the format of 2nd index CV/1st index CV.

	2nd Biomass Index Trend		
Measurement error CV	0%	-10%	+10%
LY			
CV=0.05/0.10	10.7%	9.6%	9.7%
CV=0.10/0.25	15.8%	16.6%	15.5%
CV=0.25/0.50	41.0%	70.5%	40.7%
LY/FY			
CV=0.05/0.10	10.5%	9.2%	9.2%
CV=0.10/0.25	15.3%	16.2%	15.0%
CV=0.25/0.50	40.1%	75.4%	39.8%
M			
CV=0.05/0.10	1.9%	2.7%	3.5%
CV=0.10/0.25	1.4%	1.3%	1.4%
CV=0.25/0.50	6.8%	6.1%	7.2%
q			
CV=0.05/0.10	8.7%	7.8%	8.7%
CV=0.10/0.25	11.4%	12.1%	12.0%
CV=0.25/0.50	54.5%	55.1%	55.0%
$F_{40\%}$			
CV=0.05/0.10	2.7%	3.3%	2.5%
CV=0.10/0.25	2.1%	2.4%	1.7%
CV=0.25/0.50	5.2%	5.4%	5.8%
σ_r			
CV=0.05/0.10	3.1%	3.3%	2.8%
CV=0.10/0.25	3.2%	3.7%	2.9%
CV=0.25/0.50	26.6%	23.5%	27.3%

Table 5.8. Experiment 3 relative estimation accuracy (REA) measured in relative median error for examining a second biomass index. *LY*=last year biomass, *LY/FY*=last year biomass/first year biomass, *M*=natural mortality, *q*=catchability, $F_{40\%}$=harvest rate, σ_r=recruitment variation. CVs are in the format of 2nd index CV/1st index CV.

	2nd Biomass Index Trend		
Measurement error CV	0%	-10%	+10%
LY			
CV=0.05/0.10	2.5%	-9.8%	3%
CV=0.10/0.25	-3.9%	-15.3%	0.8%
CV=0.25/0.50	-15.6%	-27.1%	-10.1%
LY/FY			
CV=0.05/0.10	2.6%	-9%	2%
CV=0.10/0.25	-3.6%	-14.0%	0.6%
CV=0.25/0.50	-14.5%	-25.5%	-9.4%
M			
CV=0.05/0.10	-1.8%	-2.9%	0%
CV=0.10/0.25	-1.8%	-3.5%	-0.7%
CV=0.25/0.50	-3.9%	-5.3%	-2.9%
q			
CV=0.05/0.10	-1.7%	5.8%	0%
CV=0.10/0.25	-6.4%	0.0%	-7.9%
CV=0.25/0.50	-36.7%	-32.4%	-38.9%
$F_{40\%}$			
CV=0.05/0.10	-1.2%	-2.3%	1%
CV=0.10/0.25	-1.2%	-2.7%	-0.1%
CV=0.25/0.50	-2.8%	-4.2%	-2.0%
σ_r			
CV=0.05/0.10	-4.2%	-0.4%	-4%
CV=0.10/0.25	-3.9%	-0.7%	-3.8%
CV=0.25/0.50	0.1%	2.4%	-0.3%

Table 5.9. Experiment 3 average estimated standard deviations and CVs of last year's total biomass from the Hessian matrix compared with the value estimated from the generating model. *LY*=last year biomass, *LY/FY*=last year biomass/first year biomass, *M*=natural mortality, *q*=catchability, $F_{40\%}$=harvest rate, σ_r=recruitment variation. CVs are in the format of 2nd index CV/1st index CV.

	2nd Biomass Index Trend		
Measurement error CV	0%	-10%	+10%
True	84211	Estimated standard deviations	
CV=0.05/0.10	61728	92165	62983
CV=0.10/0.25	80360	107005	74955
CV=0.25/0.50	79396	84462	92777
True	0.27	Estimated coefficients of variation	
CV=0.05/0.10	0.21	0.36	0.22
CV=0.10/0.25	0.30	0.45	0.26
CV=0.25/0.50	0.38	0.44	0.41

Table 5.10. Experiment 4 relative estimation accuracy (REA) measured in relative median error examining data weightings. *LY*=last year biomass, *LY/FY*=last year biomass/first year biomass, *M*=natural mortality, *q*=catchability, $F_{40\%}$=harvest rate, σ_r=recruitment variation.

	Data Source				
Relative weight	Fishery lengths	Survey ages	Fishery ages	Survey biomass	Catch
LY					
0.1	73%	10%	10%	3%	7%
0.5	19%	5%	9%	8%	7%
1	4%	4%	4%	4%	4%
2	0%	8%	5%	5%	7%
10	-16%	14%	3%	1%	7%
LY/FY					
0.2	64%	10%	7%	3%	9%
0.5	17%	3%	8%	8%	7%
1	3%	3%	3%	3%	3%
2	1%	8%	6%	5%	6%
10	-15%	13%	9%	0%	6%
M					
0.2	-10%	3%	1%	-1%	-1%
0.5	-3%	2%	1%	0%	-1%
1	1%	1%	1%	1%	1%
2	2%	-1%	-2%	0%	0%
10	7%	-6%	-4%	1%	-1%
q					
0.2	-34%	-8%	-9%	-20%	-6%
0.5	-13%	-5%	-9%	-8%	-6%
1	-4%	-4%	-4%	-4%	-4%
2	-1%	-6%	-3%	-4%	-6%
10	13%	-1%	1%	-1%	-6%
$F_{40\%}$					
0.2	-10%	-1%	12%	-2%	-1%
0.5	-6%	2%	2%	-1%	-1%
1	6%	6%	6%	6%	6%
2	18%	4%	-2%	-1%	-1%
10	53%	40%	-1%	6%	0%
σ_r					
0.2	-24%	-11%	-12%	0%	-8%
0.5	-10%	-4%	-4%	-1%	-7%
1	1%	1%	1%	1%	1%
2	2%	7%	-2%	-1%	-1%
10	42%	28%	19%	1%	-6%

Table 5.11. Experiment 5 relative estimation accuracy (REA) measured in relative median error examining prior distributions. *LY*=last year biomass, *LY/FY*=last year biomass/first year biomass, *M*=natural mortality, *q*=catchability, $F_{40\%}$=harvest rate, σ_r=recruitment variation.

Prior Mean	Prior parameter and prior CV								
	M			*q*			σ_r		
	0.005	0.01	0.02	0.10	0.20	0.40	0.10	0.20	0.40
LY									
0.5	-40%	-36%	-28%	76%	35%	10%	8%	10%	10%
0.75	-22%	-17%	-4%	47%	19%	1%	7%	8%	9%
1	-2%	4%	27%	23%	4%	-4%	0%	4%	3%
1.33	32%	55%	119%	-2%	-8%	-11%	0%	3%	8%
2	142%	174%	232%	-12%	-13%	-17%	-1%	4%	7%
LY/FY									
0.5	-33%	-29%	-22%	68%	32%	9%	6%	10%	9%
0.75	-17%	-13%	-2%	43%	17%	1%	7%	8%	8%
1	0%	3%	20%	22%	3%	-4%	2%	3%	4%
1.33	24%	42%	82%	-2%	-9%	-10%	0%	2%	7%
2	90%	104%	124%	-11%	-13%	-16%	0%	4%	6%
M									
0.5	-53%	-49%	-38%	7%	5%	1%	2%	1%	1%
0.75	-30%	-24%	-10%	5%	2%	-1%	0%	1%	1%
1	-8%	1%	20%	1%	1%	-1%	0%	1%	1%
1.33	24%	35%	67%	0%	-1%	-2%	-1%	0%	0%
2	84%	100%	128%	-3%	-3%	-3%	-2%	-1%	0%
q									
0.5	57%	48%	36%	-44%	-26%	-8%	-7%	-8%	-8%
0.75	27%	20%	5%	-32%	-15%	-1%	-6%	-7%	-7%
1	2%	-4%	-22%	-19%	-4%	5%	0%	-4%	-3%
1.33	-25%	-36%	-56%	1%	9%	12%	-1%	-3%	-6%

Table 5.11 (continued).

Prior Mean	Prior parameter and prior CV								
	M			*q*			σ_r		
	0.005	0.01	0.02	0.10	0.20	0.40	0.10	0.20	0.40
2	-62%	-67%	-75%	14%	16%	19%	0%	-4%	-6%
$F_{40\%}$									
0.5	-46%	-38%	-30%	11%	9%	-1%	6%	-1%	-1%
0.75	-24%	-19%	-8%	3%	0%	-2%	-1%	-1%	-1%
1	-7%	6%	19%	2%	6%	-2%	0%	6%	0%
1.33	21%	20%	41%	5%	4%	-3%	5%	5%	-1%
2	52%	62%	81%	-3%	-3%	3%	-2%	-1%	-1%
σ_r									
0.5	16%	10%	13%	-10%	-1%	-1%	-4%	-5%	-5%
0.75	9%	7%	2%	-5%	-2%	0%	-1%	-3%	-4%
1	1%	1%	-4%	-7%	1%	0%	-2%	1%	-6%
1.33	-5%	-10%	-20%	2%	3%	1%	8%	3%	-2%
2	-24%	-27%	-33%	1%	1%	3%	11%	3%	-1%

5.6 References

Bence, J.R., A. Gordoa and J.E. Hightower. 1993. The influence of age-selective surveys on the reliability of stock synthesis assessments. Can. J. Fish. Aquat. Sci. 50: 827-840.

Butterworth, D.S., J.N. Ianelli and R. Hilborn. 2003. A statistical model for stock assessment of southern bluefin tuna with temporal changes in selectivity. Afr. J. mar. Sci 25: 331-361.

Coggins, L. G. and T. J. Quinn II. 1998. A simulation study of the effects of aging error and sample size on sustained yield estimates. *In* Fishery Stock Assessment Models. Edited by F. Funk, T.J. Quinn II, J. Heifetz, J.N. Ianelli, J.E. Powers, J.F. Schweigert, P.J. Sullivan, and C.-I. Shang. Alaska Sea Grant College Program, Report No. AK-SG-98-01, University of Alaska, Fairbanks. pp. 315-338.

Fournier, D., and C.P. Archibald. 1982. A general theory for analyzing catch at age data. Can. J. Fish. Aquat. Sci. 39: 1195-1207.

Fu, C., and T.J. Quinn II. 2000. Estimability of natural mortality and other population parameters in a length-based model: *Pandalus borealis* in Kachemak Bay, AK. Can. J. Fish. Aquat. Sci. 55: 2420-2432.

Hanselman, D.H., T.J. Quinn II, C. Lunsford, J. Heifetz and D.M. Clausen. 2001. Spatial implications of adaptive cluster sampling on Gulf of Alaska rockfish. In Proceedings of the 17th Lowell-Wakefield Symposium: Spatial Processes and Management of Marine Populations, pp. 303-325. Univ. Alaska Sea Grant Program, Fairbanks, AK.

Hanselman, D., J. Heifetz, J. Fujioka, and J. Ianelli. 2003a. Gulf of Alaska Pacific ocean perch. In Stock Assessment and fishery evaluation report for the groundfish resources of the Gulf of Alaska. North Pacific Fisheries Management Council, Anchorage. pp. 429-479.

Hanselman, D.H., T.J. Quinn II, C. Lunsford, J. Heifetz and D.M. Clausen. 2003b. Applications in adaptive cluster sampling of Gulf of Alaska rockfish. Fishery Bulletin 101:501-512.

Merritt, M.F. and T.J. Quinn II. 2000. Using perceptions of data accuracy and empirical weighting of information: assessment of a recreational fish population. Can. J. Fish. Aquat. Sci. 57: 1459-1469.

Methot, R.D., 1989. Synthetic estimates of historical abundance and mortality for northern anchovy. Amer. Fish. Soc. Symp. 6: 88-82.

Myers, R.A., Bridson, J. and Barrowman, N.J. 1995. Summary of worldwide spawner and recruitment data. *Can. Tech. Rep. Fish. Aquat. Sci.* No. 2024, 327 pp.

National Research Council (NRC). 1998. Improving fish stock assessments. National Academy Press, Washington, D.C. 177 pp.

Parma, A.M. 1993. Retrospective catch-at-age analysis of Pacific halibut: implications on assessment of harvest policies. *In* Proceedings of the International Symposium on Management Strategies for Exploited Fish Populations. *Edited by* G. Kruse, D.M. Eggers, R.J. Marasco, C. Pautzke, and T.J. Quinn II. Alaska Sea Grant College Program, Report No. 93-02, University of Alaska, Fairbanks. pp. 247-265.

Punt, A.E., A.D.M. Smith and G. Cui. 2002. Evaluation of management tools for Australia's south east fishery. 2. How well can management quantities be estimated? Mar. Freshwater Res. 53: 631-644.

Punt, A.E. and R. Hilborn. 1997. Fisheries stock assessment and decision analysis: the Bayesian approach. Reviews in Fish Biology and Fisheries 7: 35-63.

Quinn, T.J., II, and R.B. Deriso. 1999. Quantitative fish dynamics. Oxford University Press, New York.

Sampson, D.B. and Y. Yin. 1998. A Monte Carlo evaluation of the stock synthesis assessment program. *In* Fishery Stock Assessment Models. *Edited by* F. Funk, T.J. Quinn II, J. Heifetz, J.N. Ianelli, J.E. Powers, J.F. Schweigert, P.J. Sullivan, and C.-I. Shang. Alaska Sea Grant College Program, Report No. AK-SG-98-01, University of Alaska, Fairbanks. pp. 315-338.

General Conclusions

Stock assessment is a science based in both mathematics and art. It involves choices on using data and models with many inherent problems: (1) There is never enough data. (2) The data are rife with error and uncertainty. (3) The mathematics are only a guide; the practitioner must make decisions about the biological realism of model results before presenting to managers. (4) The survey is never good enough; stock assessment biologists always want more samples and more precision. These are just some of the problems in dealing with stock assessment and surveys.

The assessment of Pacific ocean perch is an excellent case study for these issues. In Chapter 1, I presented the stock assessment of Pacific ocean perch. In the previous year's stock assessment (Heifetz et al. 2002), I explored the use of Markov Chain Monte Carlo to examine uncertainty in the model. This resulted in uncovering some model misspecification problems. In Chapter 1, I implemented new length to age matrices that included density dependent growth that greatly improved the fit of the model to the data. I also showed which parameters are uncertain and gave a more realistic portrayal of the uncertainty in the biomass estimates. After resolving some of these issues in the stock assessment model, the survey biomass estimates have more relative importance as a data source.

In Chapter 2, I examined the use of the adaptive cluster sampling as an alternative design to improve survey precision. This was the second of two studies using the design. This field test provided better results than the first attempt at the design (Hanselman et al. 2001). The estimates of mean abundance had lower standard errors than simple random sampling, although the mean was much lower for the second time in two studies. I showed by reconstructing adaptive samples with a higher criterion value, that similar precision could be gained with a smaller sample size. The higher criterion value would enable the survey to utilize a cost-recovery charter, because the adaptive sampling would concentrate only on high density areas. Bootstrap results suggested that adaptive sampling may normalize the skewed population distribution of Pacific ocean perch and its estimate of the mean may be a more reliable measure of central tendency. Although the gains in precision were not as good as expected, the technique showed some potential for gaining more samples quickly and inexpensively, especially if a cost-recovery charter conducts the additional sampling.

In Chapter 3, I use simulations to examine when adaptive cluster sampling might work better than it did in previous field tests. When sampling a population constructed with the characteristics of Pacific ocean perch, adaptive sampling did not become more efficient than simple random sampling until more than 10% of the sampling frame was enumerated. If such sampling effort could be applied, adaptive sampling becomes much more efficient. Unfortunately, in large marine populations this type of coverage is unlikely. This conclusion may be a general one. In Su and Quinn (2003), a highly aggregated population 25% the size of the one simulated for this study was sampled. Those results also showed that large gains in efficiency were not attained until a large proportion of the population had been sampled.

In Chapter 4, I explored the use of hydroacoustic data gathered during summer surveys in the Gulf of Alaska, Aleutian Islands, and Bering Sea. Pacific ocean perch and other rockfish have prominent swim bladders which make them a candidate for hydroacoustic assessment, but live near the bottom and may be difficult to detect (Stanley et al. 1999). I used the echointegrated acoustic backscatter from the area corresponding to the height of the survey net for sections that were trawled, and sections that were not. I used these data in a model that makes predictions for tows within 10 km of actual tows. The predictions were used for adaptive sampling, double sampling, and to pool strata from the 2001 groundfish survey. All three uses of the hydroacoustic data showed gains in precision of biomass estimates, but double sampling was probably the most theoretically correct use of the auxiliary information. In double sampling, the auxiliary variable can be "eyeball estimates" or model predictions. If this information is highly correlated with the variable of interest then significant precision can be gained with negligible bias (Thompson 2002). The main problem with this method was that the estimation was done on the transformed scale to obtain this high correlation. When this estimate was back-transformed, it yielded a geometric mean (Taylor 1986) which made it hard to interpret as a measure of absolute biomass but could be useful as a secondary index.

In Chapter 5, I conducted simulations to examine properties of the Integrated Analysis type of stock assessment used for many species on the west coast of North America. I found that species with a high level of measurement error in the survey biomass estimates may be subject to significant biases in the stock assessment. Trends in catchability are more important in a precise survey than an imprecise survey because the model predicts the opposite trend in catchability at high survey precision and is greatly biased by the trend. The addition of an additional survey index reduced the uncertainty in the results of the model, and gave more total weight to the average trend of the two indices, hence reducing influence of other data sources. If other data sources are unreliable, additional survey biomass estimates will reduce influence from these other sources. Data weighting was less influential than expected. The only data source that had a large impact on biomass estimates and other key parameters was the fishery length data. This is the largest data set in the model, and the only one other than catch that has data prior to 1984. Fishery length data had the largest influence on estimated total biomass in the beginning of the time series, therefore greatly affected all future trajectories. The choice of prior distributions was important for two of the three parameters that utilize priors in the model. Natural mortality was extremely sensitive to both the precision of the prior and its mean, with small perturbations resulting in large changes in estimated biomass. Model results were sensitive to catchability priors when they were assumed to be known precisely, but were insensitive when the prior was specified imprecisely.

Unlike terrestrial stock assessment, assessment of deep-water marine stocks is always a situation of incomplete information. Barring an evolutionary leap in survey methodology, this situation will continue far into the future. In the mean time, stock assessment biologists must continue to make tradeoffs on what data is the most important, reliable, and cost-effective. This thesis mainly explored survey biomass estimates and their relation to the stock assessment model. These models contain a plethora of other uncertainties to be explored, and research must be continued in an effort to optimize and validate their use.

References

Hanselman, D.H., T.J. Quinn II, C. Lunsford, J. Heifetz and D.M. Clausen. 2001. Spatial inferences of adaptive cluster sampling on Gulf of Alaska rockfish. Proceedings of the 17th Lowell-Wakefield Symposium: Spatial Processes and Management of Marine Populations, 303-325.

Heifetz, J., D.L. Courtney, D. M. Clausen, D. Hanselman, J. T. Fujioka, and J. N. Ianelli. 2002. Slope rockfish. In Stock assessment and fishery evaluation report for the groundfish resources of the Gulf of Alaska as projected for 2002. North Pacific Fishery Management Council, 605 W 4th Ave, Suite 306 Anchorage, AK 99501. pp.295-382.

Stanley, R. D., R. Kieser, B.M. Leaman and K. Cooke. 1999. Diel vertical migration by yellowtail rockfish (*Sebastes flavidus*) and its impact on acoustic biomass estimation. Fishery Bulletin 97: 320–331.

Su, Z. and T. J. Quinn. 2003. Estimator bias and efficiency for adaptive cluster sampling with order statistics and a stopping rule. Environmental and Ecological Statistics 10:17-41.

Taylor, J.M.G. 1986. The retransformed mean after a fitted power transformation. Journal of the American Statistical Association 81:114-118.

Thompson, S.K. 2002. Sampling. 2nd edition, Wiley, New York.